Nazarova L, Self AE, Brooks SJ, van Hardenbroek M, Herzschuh U, Diekmann B. Northern Russian chironomid-based modern summer temperature data set and inference models. *Global and Planetary Change* 2015, 134, 10-25.

Northern Russian chironomid-based modern summer temperature data set and inference models

Larisa Nazarova[a,b*], Angela E. Self[c], Stephen J. Brooks [c], Maarten van Hardenbroek[d], Ulrike Herzschuh[a], Bernhard Diekmann[a]

[a] *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Research Unit Potsdam, Telegrafenberg A43, 14473 Potsdam, Germany*

[b] *Kazan Federal University, Kremlyovskaya str., 18, 420018 Kazan, Russia*

[c] *Department of Life Sciences, Natural History Museum, Cromwell Road, London SW7 5BD, UK*

[d] *Geography and Environment Department, University of Southampton, University Road, Southampton SO17 1BJ, UK*

*Corresponding author: Tel: +49 30 31800549; Fax: +49 30 31800549; E-mail: nazarova_larisa@mail.ru

E-mail addresses: *larisa.nazarova@awi.de (L. Nazarova), a.self@nhm.ac.uk (A.E. Self), s.brooks@nhm.ac.uk (S.J. Brooks), vanhardenbroek@soton.ac.uk (M. van Hardenbroek), Ulrike.herzschuh@awi.de (U. Herzschuh), bernhard.diekmann@awi.de (B. Diekmann)*

Key words

Chironomidae, arctic Russia, transfer function, ecology, biogeography, temperature

Abstract

West and East Siberian data sets and 55 new sites were merged based on the high taxonomic similarity, and the strong relationship between mean July air temperature and the distribution of chironomid taxa in both data sets compared with other environmental parameters. Multivariate statistical analysis of chironomid and environmental data from the combined data set consisting of 268 lakes, located in northern Russia, suggests that mean July air temperature explains the greatest amount of variance in chironomid distribution compared with other measured variables (latitude, longitude, altitude, water depth, lake surface area, pH, conductivity, mean January air temperature, mean July air temperature, and continentality). We established two robust inference models to reconstruct mean summer air temperatures from subfossil chironomids based on ecological and geographical approaches. The North Russian 2-component WA-PLS model ($RMSEP_{Jack}$ = 1.35 ^{o}C, r^2_{Jack} = 0.87) can be recommended for application in palaeoclimatic studies in northern Russia. Based on distinctive chironomid fauna and climatic regimes of Kamchatka the Far East 2-component WAPLS model ($RMSEP_{Jack}$ = 1.3 ^{o}C, r^2_{Jack} = 0.81) has potentially better applicability in Kamchatka.

1. Introduction

Biotic proxies from lake sediments provide a powerful means of quantifying past climate change in terrestrial contexts. In addition, analysis of biotic remains from lake sediments provides an indication of the rate and magnitude of the response of animals and plants to past climate change and how they may respond in the future. Climatic inferences from palaeorecords are based on modern or near-modern analogues (training sets) from which the empirical reconstruction models (i.e. the transfer function) are established. By using inference models, which link the present distribution and abundance of chironomids to contemporary climate, past climates can be quantified from fossil chironomid assemblages (Self et al., 2011; Kienast et al., 2011). Chironomids (Insecta: Diptera) are well-proven to be among the most reliable quantitative proxies of mean July air temperature (Brooks, 2006). They are a diverse and nearly ubiquitous family of holometabolous two-winged flies and play a vital role in freshwater ecosystems as primary consumers (Coffman and Ferrington, 1996). The abundance and distribution of most chironomid taxa are temperature-dependent (Walker et al., 1991), reflecting the effect of air and water temperatures on all stages of their life cycles (Oliver, 1971) and they respond rapidly to climate change by virtue of the winged adult stage. The larval head capsules preserve well in lake sediment deposits and the subfossils are readily identifiable in most cases at least to species morphotype (Brooks et al., 2007).

Chironomid based inference models for reconstructing mean July air temperature have been developed successfully for Western Europe (Olander et al., 1999; Brooks and Birks, 2001), North America (Walker et al., 1997; Barley et al., 2006), Africa

(Eggermont et al., 2007), New Zealand (Woodward and Shulmeister, 2006) and Tasmania (Rees et al., 2008).

Recently, data on the distribution and abundance of chironomids in lakes along environmental gradients in eastern and western Siberia were used to develop modern chironomid-based calibration data sets (training sets) and quantitative transfer functions for reconstructing mean July air temperature (T July), water depth (WD) and continentality (CI) in eastern (ES) and western Siberia (WS) (Nazarova et al., 2011; Self et al., 2011). Numerical analysis showed that T July is the most significant variable explaining contemporary chironomid distribution and abundance in both data sets. These data sets and transfer functions have provided a new tool for quantitative assessment of the past environment in north-eastern Eurasia and were applied in several studies of Holocene palaeoclimate in Siberia (Kienast et al., 2011; Self et al., 2011; Jones et al., 2011; Mackay et al. 2012; Nazarova, 2012, 2013a, 2013b; Engels et al., 2014). Climate inference models have limited application outside the regions in which they have been developed, so the new Russian models are an improvement over the Swedish and Norwegian inference models, which have been used previously for chironomid-inferred temperature reconstructions in northern Russia. Solovieva et al. (2005) reconstructed T July in north-east European Russia using a chironomid July air temperature-inference model based on a modern training set of 153 Norwegian lakes (Brooks and Birks, 2001 and unpublished data), supplemented with data from lakes within the study area. The chironomid temperature-inference model developed for northern Sweden (Larocque et al., 2001) has also been used for temperature reconstructions in the Lena River Delta

(Andreev et al., 2004), the Kola Peninsula (Ilyashuk et al., 2005) and Polar Urals (Solovieva et al., 2005).

In this paper we present the results of the work we have done to re-analyse and standardise the taxonomy between our already published chironomid calibration sets from East Siberia (Nazarova et al., 2011) and West Siberia (Self et al., 2011) with the addition of new regions to the data set: Bunge Land (Laptev Sea), 31 lakes from Kolyma River region, 10 lakes from Indigirka River region and 13 lakes from Kamchatka (Fig. 1). Following taxonomic standardisation we have merged the data sets. This has the advantage of extending the geographical and environmental gradients and increasing the representation of taxa in the calibration set. This can be expected to further improve the performance and applicability of the chironomid-temperature inference models by providing better estimates of the environmental optimum of taxa and increasing the probability of analogues between present and past assemblages.

The main objectives of our investigation are to compare the faunal composition of the WS and ES data sets, to examine the environmental factors which influence chironomid distribution and abundance in the combined data set, to identify which climate variable has the most potential for development of a chironomid-based inference model and to develop chironomid inference models for quantifying past regional climate and environmental changes in northern and north-eastern Russia.

2. Study regions

Study sites included in this investigation span wide latitudinal and longitudinal ranges in northern Russia: from Komi Republic in the West (50.50 E, part of the WS) to

Kamchatka in the East (163.15 E, new data) and from Novosibirsk Islands in the Laptev Sea in the North (75.40 N, part of the ES) to southern Kamchatka (53.03 N, new data) in the South (Fig. 1).

The most western part of the data set includes the Komi Republic (region Komi, part of WS, Fig. 1) and Bolshezemelskaya tundra (region Pechora, part of WS, Fig. 1), a large lowland plain situated to the west of the Ural Mountains within the zone of discontinuous - continuous permafrost (Fig. 1). Climate is severe with an eight-month winter period when mean monthly temperatures are below 0 ^{o}C. Mean July air temperatures are about 12.8 ^{o}C (New et al., 2002). Annual precipitation varies between 370 and 395 mm (Mukhin et al., 1964; Solovieva et al., 2005, 2008).

The next region, east, included in the data set is the Putorana Plateau, a mountainous area at the north-western edge of the Central Siberian Plateau and to the south of the Taymyr Peninsula (region Putoran, part of WS, Fig. 1). The climate is continental with short warm summers and long, cold winters. Most of the territory is covered with conifer forests. Investigated lakes here crossed the latitudinal tree line from sparse larch forest to alpine tundra. The coldest month is February with mean temperatures of about -32.0 ^{o}C and the warmest month is July with mean temperatures of 11.8 ^{o}C (New et al., 2002).

Most of the lakes in the data set were collected from different regions of Yakutia (regions Anabar, Lena, Laptev, Central Yakutia (CY), Northern Yakutia (NY) are parts of WS and ES, regions Indigirka, Kolyma are newly investigated, Fig. 1) (between 55.48 and 76.77 N and 105.53 and 162.92 E; Fig. 1). This part of Eurasia has one of the Earth's most extreme semi-arid continental settings (Kumke et al., 2007). It is characterized by pronounced seasonality. The coldest month is January, with average temperatures below -

40°C (New et al., 2002) and minimum temperatures less than -60.0°C (Gavrilova, 1998). Average July temperature varies from about 2.0–4.0°C on the New Siberian Islands in the Laptev Sea (region Laptev, part of ES, Fig. 1), to about 18.0–19.0 °C in CY near Yakutsk (Fig. 1) (New et al., 2002). Annual precipitation ranges between 141 and 546 mm, which is less than annual evaporation in most areas (Gavrilova, 1998). The driest area is the Central Yakutian lowland, where summer evaporation is four times higher than precipitation.

The most eastern region in the data set is Kamchatka, a 1250-kilometre peninsula in the Russian Far East that lies between the Pacific Ocean to the east and the Sea of Okhotsk to the west (region Kamchatka, newly investigated, Fig. 1). The geographic position, the surrounding seas and the high relief are all factors contributing to diverse and variable weather in Kamchatka. The highest monthly mean temperature (July) ranges from 12.0 to 16.0 °C and the lowest monthly mean temperature (January) ranges from -20.0 °C to -18.0 °C (New et al., 2002). Annual precipitation in the Central Valley is 350 mm (Krestov et al., 2008; Neshataeva, 2008).

The geographical area included in the study covers major biomes, such as the boundaries between continuous and discontinuous permafrost zones and from boreal coniferous forests in the south to tundra vegetation in the north. The altitude of the sampled lakes varies from 5 m below sea level (in Kolyma region, North-Eastern Yakutia, Fig. 1) to 1218 m above sea level (a.s.l.) (in Kamchatka). The mean July air temperature ranges from 1.8 to 18.9 °C, the mean January air temperature (T Jan) ranges from -45.3 to -11.2 °C (New et al., 2002), annual precipitation lies between 150 mm in CY to 1350 mm in Kamchatka (Gavrilova, 1998).

Quaternary deposits are widely distributed throughout the study regions. The lakes in the Komi Republic are formed on a sedimentary sequence of Palaeozoic carbonate and deep marine shale (Lindquist, 1999). The remaining lakes from the Putorana Plateau, the southern Lena Delta and near Vilyuysk, in CY, lie on the Siberian Traps, a large igneous province extruded in the late Permian ca 240-220 Myrs (Zolotukhin and Almukhamedov, 1988). The geological subsoil (a layer of earth immediately below the surface soil), of the lakes in Kamchatka is determined by Plio-Pleistocene volcanic rocks (IFM-GEOMAR 2008).

3. Methods

3.1. Field methods and derivation of climate variables

Surface sediments and environmental data for the West Siberian (WS) data set (100 lakes in total; Self et al., 2011) were collected between 1998 and 2007 in north-east European Russia (Solovieva et al., 2002, 2005; Sarmaja-Korjonen et al., 2003), Lower Lena River (Porinchu and Cwynar, 2000), in Putorana Plateau, in the Komi Republic (Self et al., 2011). The sediment samples and environmental data for the East Siberian (ES; Nazarova et al., 2005, 2008, 2011) data set, collected from 2003–2007, includes 150 lakes from several regions of Yakutia (Fig. 1). The lakes span wide latitudinal and longitudinal ranges and are distributed through several environmental zones (arctic tundra, typical tundra, steppe-tundra, boreal coniferous forest), but all are situated within the zone of continuous permafrost. Most of the sampled lakes are small and shallow with little or no inflow. Lakes have a simple morphology and a single basin.

We included in the present study 55 new lakes: one unnamed lake from Bunge Land (75.40 N 141.26 E, Novosibirsk Islands, Yakutia, region Laptev), 31 lakes from the Kolyma River region, North-Eastern Yakutia (68.22 to 69.33 N, 159.97 to 161.88 E, region Kolyma, Fig. 1) collected in 2011, 5 lakes from central and southern Kamchatka, collected in 2009, 8 lakes from northern and central Kamchatka collected in 2005 (region Kamchatka, Fig. 1), 10 lakes from Indigirka River region in North-Eastern Yakutia (70.74 to 70.82 N, 147.42 to 147.59 E, Indigirka, see details in van Hardenbroek et al., 2013, Fig. 1) collected in 2007.

The sampling methods are broadly similar throughout the study. Sediment cores were collected from the deepest point of each lake using an 80 mm diameter HON-Kajak corer (Renberg, 1991) or 60 mm diameter UWITEC gravity corer. Cores were extruded in the field at 0.25-1.0 cm intervals. Samples were stored in whirl-pak bags and kept cool and dark in the field prior to storage at 4.0 ^{o}C. The water depths (WD) were measured using an echolot. Total hardness, alkalinity, and acidity of East Siberian samples, collected by AWI, were determined using titrimetric test kits (Aquamerck in 2002, Macherey-Nagel, Visocolor series in 2005 and 2007). Water temperature, pH and electrical conductivity were quantified using a handheld multi-parameter instrument (WTW 330i, 340i). More details on the treatment of water samples from West Siberia and Indigirka can be found in Self et al (2011) and van Hardenbroek et al (2013) respectively.

T July and mean January air temperatures (T Jan) for each site were obtained from a climatic data set compiled by New et al. (2002), measured at 2 m above the ground in standard meteorological screens. Using these data, the mean air temperatures at each lake were estimated by spatial interpolation of elevation and distance from the coast. The New

et al. (2002) data set uses climate normals from 1961 to 1990 to create a global climatic grid with a resolution of 10-min latitude/longitude. Some shortfalls are associated with this data set, for example, the relatively coarse resolution of the climate data and the climate normals predate the collection of chironomid data and span a cold phase of the Arctic Oscillation (Overland and Wang, 2005). However the global nature of the New et al. (2002) data set, its homogeneity and consistency make it more suitable for our purpose in comparison to the data that was obtained from local meteorological stations. These latter data have gaps in observations and cover different time spans.

Gorczynski`s continentality Index (1920) modified by Grieser et al. (2006) was used in this study as it is easily computed and applicable to high latitudes (Self et al., 2011):

$$CI = 1.7(A/\sin \varphi) - 20.4$$

where A is the annual range of average monthly temperatures in ^{o}C and φ the latitude of the lake. The annual temperature range was based on T July and T Jan calculated from New et al. (2002).

3.2. Chironomid analysis

Treatment of sediment samples for chironomid analysis followed standard techniques described in Brooks and Birks (2000). Subsamples of wet sediments were deflocculated in 10 % KOH, heated to 70 ^{o}C for up to 10 minutes and left for another 20 minutes after adding boiling water. The sediment was then passed through stacked 212 and 90 µm sieves. Chironomid larval head capsules were picked out of a grooved Bogorov sorting tray using fine forceps under a stereomicroscope at 25–40x magnification. Larval head capsules were mounted two at a time in Euparal or Hydromatrix, ventral side up, under a 6 mm diameter cover slip, with ten cover slips per

microscope slide. In total, 26 221 chironomid head capsules were slide mounted and identified.

Prior to merging the datasets, to ensure taxonomic harmonisation, chironomids were re-identified and re-analysed to the highest taxonomic resolution possible with reference to Wiederholm (1983), Makarchenko and Makarchenko (1999), Brooks et al. (2007) and at several taxonomic workshops between the authors of this paper.

Between 50 and 560 head capsules were identified from the majority of surface sediments. However only 44-49 head capsules were isolated from 13 high-latitude lakes (Lena, Anabar) and 20 to 40 head capsules were isolated from eight further high-latitude lakes (Kolyma). Nevertheless, the typically low taxonomic diversity of these assemblages suggests these abundances are acceptable for inclusion in the training set (Quinlan and Smol, 2001a).

3.3. Water chemistry analysis

Analyses of water samples from the East Siberian data is presented in Nazarova et al. (2011). Data from north-east European Russia (Komi, Pechora) and Putorana Plateau (Fig. 1) are presented in Self et al. (2011). Analyses of water samples from the Lena Delta are described by Duff et al. (1998).

Water samples of 31 lakes from Kolyma region and of 5 lakes from Kamchatka were analysed at AWI. Samples were initially passed through a cellulose-acetate filter (pore size 0.45 µm) in the field. Afterwards, samples for cation analyses were acidified with HNO_3 whereas samples for anion analysis and residue samples were stored in thermoboxes to ensure cool conditions. Upon return to the laboratory, the cation content

of the water was analysed by Inductively Coupled Plasma-Optical Emission Spectrometry (ICP-OES, Perkin-Elmer Optima 3000 XL) while the anion content was determined by Ion Chromatography (IC, Dionex DX-320). The bicarbonate concentrations of the water were calculated from the alkalinity measurements in the field and additionally checked by titration with 0.01 M HCl using an automatic titrator (Metrohm 794 Basic Titrino). In each lake, 25–30 different chemical and environmental variables were measured.

For the water samples from 8 further Kamchatka lakes and from 10 lakes from Indigirka region only a limited number of chemical and environmental data are available for both of these two regions: latitude (N), longitude (E), altitude (m a.s.l.), Water depth (WD, m), Lake area (ha), pH, conductivity (µS/cm), T Jan (°C), T July (°C) and CI. Lake area data are not available for 92 lakes (North-Eastern Yakutia, Nazarova et al. 2011; region NY, Fig. 1), therefore, this parameter was excluded from further analysis as well.

Thus, nine environmental parameters were available for all lakes in the present investigation and were used in statistical analyses. A summary of the environmental variables in the 268 lakes data set is presented in Table 1.

3.4. Numerical methods

The combined 268 lakes data set was analysed to examine the relationship between the environmental variables and chironomid distribution and abundance. All taxon data were transformed to percent abundances, calculated as percentage of total identifiable chironomids (Brooks and Birks, 2001; Barley et al., 2006) and were square root transformed prior to analysis. Environmental variables with skewed distributions (Altitude, WD and conductivity) were log transformed. Skewness reflects the degree of

asymmetry of a distribution around its mean. Normal distributions produce a skewness statistic of about zero. Values that exceeded 2 standard errors of skewness (regardless of sign) were identified as significantly skewed (Sokal and Rohlf, 1995). Remaining parameters were left untransformed.

Detrended Correspondence Analysis (DCA) with detrending by segments was performed on the chironomid data (rare taxa downweighted) to explore the main pattern of taxonomic variation among sites and to determine the lengths of the sampled environmental gradients, from which we decided whether unimodal or linear statistical techniques would be the most appropriate for the data analysis (Birks, 1995). The gradient length of species scores was relatively long. DCA axes 1 and 2 were 3.78 and 4.12 standard deviation units respectively, indicating that numerical methods based on a unimodal response model were the most appropriate to assess the variation structure of the chironomid assemblages (ter Braak, 1995).

Variance inflation factors (VIF) were used to identify intercorrelated variables. Environmental variables with a VIF greater than 20 were eliminated, beginning with the variable with the largest inflation factor, until all remaining variables had values < 20 (ter Braak and Šmilauer, 2002b).

Relationships between chironomid distribution and environmental variables were assessed using a set of Canonical Correspondence Analyses (CCA) with each environmental variable as the sole constraining variable. The percentage of the variance explained by each variable was calculated. Statistical significance of each variable was tested by a Monte Carlo permutation test with 999 unrestricted permutations (ter Braak,

1990). Significant variables ($P \leq 0.05$) were retained for further analysis. Both DCA and CCA were performed using CANOCO 4.5 (ter Braak and Šmilauer, 2002a). Taxonomic similarities between regions of investigation within the data set was estimated using DCA, Sørensen–Dice index of taxonomic similarity (TSI) and squared-chord distance as measure of taxonomic distances (TD) (Gavin et al., 2003) performed in the program PAST (Hammer et al., 2001). TSI ranges from 0 to 1, with a score of 0 indicating that sites have an identical taxonomic composition and a score of 1 that sites are completely different (Sørensen, 1957). Squared chord distance values can range from 0.0 to 2.0, with 0.0 indicating identical proportions of species within the samples being compared (Dowsett and Robinson, 1998).

3.5. Model development

The environmental variable explaining most variance in the data set (indicated by the CCAs) was used to develop quantitative transfer functions based on weighted averaging (WA) and weighted averaging partial least squares (WA-PLS) methods (ter Braak and Looman, 1986; ter Braak and Juggins, 1993; Barley at al., 2006). Relationships between the significant environmental variables and the individual axes in the models are examined through correlation coefficients, t-values and interset correlations. The critical value for a t-test is 2.1, at the 5% significance level (ter Braak and Šmilauer 2002b).

The performance of the models and optimal number of components in the transfer function were assessed using leave-one-out, jack-knifed, cross validation. Each inference model was evaluated by means of the coefficient of determination (r^2_{Jack}), root mean squared error of prediction (RMSEP), a measure of random error in the model (Altman

and Bland, 1983) and max bias $_{Jack}$ (the tendency of the model to over- or underestimate along a particular portion of the gradient). The most robust transfer functions were those that had the lowest RMSEP, high coefficient of determination (r^2_{Jack}) and low mean and max bias $_{Jack}$. The number of components included in the final model was selected based on reducing the RMSEP by at least 5% (Birks, 1998).

Lakes were defined as outliers based on their ecological parameters (Behar, 1997) or if the absolute residual of the samples exceeded the standard deviation of T July in all trial models (Birks et al., 1990a). Five of the thirteen lakes from Kamchatka appeared to be outliers in NR model. Exclusion of these lakes did not influence statistical parameters of the obtained model and these were left in the model for better representation of the Kamchatkan distinctive fauna.

In order to improve statistical parameters of the models we used ecological and geographical approaches to lake selection. The lakes in the data set are unevenly distributed along the T July gradient, as 66 lakes located in CY have T July of 16.0-19.0 ^{o}C. In order to minimise variation in the other variables we deleted 57 lakes with conductivity above 500 µS/cm and with pH ≥8 (Behar, 1997). In addition to the lakes with high conductivity and pH, we deleted 18 other outliers, based on their high residuals in the model. In total we excluded 75 lakes, mostly from CY (high conductivity, high pH), Laptev (high pH, WD≤0.3 m) and Kolyma River basin (high conductivity, high pH).

In order to produce a T July model that can have a potentially better applicability in the Russian Far East we used a geographical approach for selecting lakes for the Far East (FE) data set. We ranked all the lakes by longitude. There was a natural break in the data at approximately 140^{o}E (between 133^{o} and 141^{o}E) which corresponds to the Verkhoyansk

Mountains. The Far East training set contained only sites east of longitude 140°E, including 102 lakes from Kamchatka, Kolyma, NY, Indigirka and Laptev regions. Fourteen lakes with high conductivity or pH appeared to be outliers and were deleted from the data set.

Optimum and tolerance for all chironomid taxa retained in the analyses were estimated by weighted averages and weighted standard deviations (Birks et al., 1990b). Optima of taxa with N2 values of 5 or more in the modern data are likely to be reliably estimated whereas optima for taxa with N2 less than 5 are likely to be poorly estimated because these taxa are rare in the modern data (Brooks and Birks, 2001). The program C2 version 1.5 (Juggins, 2007) was used to develop and assess transfer functions and estimate optima and tolerances. In order to describe the relationship of each taxon to the main environmental variables, taxon response models were generated using generalized linear models (GLM), set to a quadratic degree and Poisson distribution ($p<0.05$, significant; highly significant, $p<0.001$), in the CanoDraw component of CANOCO 4.5 (ter Braak and Šmilauer, 2002b). Taxon responses to July air temperature in the WS data set were examined by fitting response models to species abundance and presence/absence data for non-rare taxa (with 10 or more occurrences) using Huisman, Olff and Fresco (HOF) models (Oksanen and Minchin, 2002) and GLR models (Self et al., 2011). In the ES data set estimation of response to T July was based on GLR models (Nazarova et al., 2011).

4. Results

4.1. Distribution of the lakes in the Russian training sets along the T July gradient

The distribution of the sampled lakes along the T July gradient is shown in Fig. 2. The West Siberian data set (WS) covers the T July gradient from 10.8 to 18.3 °C with a high proportion of lakes between 11.5-12.0°C and very few lakes in the range between 15.0-16.4°C (Fig. 2). The East Siberian data set (ES) covers a longer T July gradient, from 3.4 to 18.8 °C, but has two major gaps in the sampled gradient. There are no lakes from 4.0 to 9.0 °C and from 13.5 to 16.6 °C. Merging the two data sets (WS and ES) and adding new 55 lakes significantly improves the distribution of the sampled lakes along the temperature gradient. The gap between 4.0 to 9.0 °C is filled with the newly investigated lakes from the Kolyma region in North-Eastern Yakutia and lakes in Kamchatka. Lakes from the Bolshezemelskaya tundra, in the Pechora region, fill in the gap between 13.5 and 16.6 °C.

The full combined data set (FM) includes 66 lakes located in CY with air temperatures of 16.0-19.0 °C. Many of these lakes have high conductivity and 16 of them have conductivities between 1128 and 7744 µS/cm (see ESM 1in Nazarova et al., 2011) and pH ≥8 due to the negative water balance in the region. A group of 72 lakes with T July between 10.5 and 11.5° C includes lakes from the Anabar region in North-Western Yakutia, the Putorana Plateau and the Lena River Delta. The 27 coldest lakes from the Novosibirsk Islands, Laptev region (T July 1.8 – 3.9 °C), are very shallow (≤0.65 m) with low conductivity (median 113 µS/cm).

4.2. Chironomid fauna

4.2.1. Taxonomic composition of newly investigated sites: Kamchatka and Kolyma

In 13 investigated lakes from Kamchatka we identified 77 chironomid taxa. Most widely distributed taxa are *Tanytarsus mendax*-type, *Procladius*, *Psectrocladius sordidellus*-

type, *Ablabesmyia*, *Limnophyes* – *Paralimnophyes* and *Paratanytarsus penicillatus*-type. *Chironomus anthracinus*-type, *Sergentia coracina*-type, *Microtendipes pedellus*-type, *Cladotanytarsus mancus*-type, *Dicrotendipes nervosus*-type are found in 7 of 13 lakes. Distribution of *Allopsectrocladius*, *Diamesa zernyi/cinerella*-type, *Eukiefferiella fittkaui*-type, *Eurycnemus*, *Psectrocladius calcaratus* type, *Pseudodiamesa* and *Rheocricotopus* is restricted to Kamchatkan lakes in our data set (not in the Fig.3).

In total, 70 taxa are identified from 31 lakes from Kolyma regions. The most abundant taxa are *C. anthracinus*-type, *Zalutschia zalutschicola, Z.* type B, *Cladopelma*, *Corynoneura*. In Kolyma region *Corynocera ambigua* has the highest abundance in the whole data set (Fig.3).

4.2.2. Taxonomic composition of the data set and inter-regional comparison

A very high TSI (0.89) was found when comparing WS and ES data sets, indicating a high taxonomic similarity. The WS data set includes 36 lakes from CY, that were also included in the ES data set and 21 lakes from the Lower Lena Delta (Porinchu and Cwynar, 2000) are geographically situated in the NY. Excluding CY lakes from the WS data set, decreased taxonomic similarity between the two data sets, but it still remained high with TSI = 0.77.

Most of the chironomid head capsules in the combined data set (96.6 %) could be identified to species, species group or to genus level and 3.4% (899 HC) to subfamily level only. In total, the combined data set includes 177 taxa within 88 genera and 4 subfamilies. Sixty-four taxa belonged to subfamily Chironominae (41 Chironomini, 22 Tanytarsini and 1 Pseudochironomini), 99 taxa to subfamily Orthocladiinae, 8 taxa to subfamily Diamesinae and 7 to Tanypodinae. Of these, 50 taxa were are and occurr in 3

or fewer lakes. The majority of these rare taxa occur in less than 5% abundances in any lake. The exceptions are *Propsilicerus jacuticus*-type with a maximum abundance of 8.8%, *Parachaetocladius* (6.2%), and *Paralauterborniella* (6.7%).

The most common taxa are *Ps. sordidellus*-type, *C. anthracinus*-type, *Procladius*, *Tanytarsus lugens*-type, *P. penicillatus*-type, *Limnophyes – Paralimnophyes* and *T. mendax*-type.

There is a high turnover of taxa across the T July gradient (Fig. 3). Sites with the coldest T July (1.8–5.0 ^{o}C) in our data set are all from the Laptev region and included Yana-Indigirka lowland, Novosibirsk Islands and the coast of the Dmitry Laptev Strait of the Laptev Sea. In total, 62 taxa are identified from this region. The most abundant taxa are *C. anthracinus*-type, *Ps. sordidellus*-type, *Orthocladius* type-I, *P. austriacus*-type, *Derotanypus, P. penicillatus*-type, and *Paratanytarsus* (HC without mandibles). *Orthocladius* type I and *P. austriacus*-type are also found in slightly warmer conditions (Fig. 3) but both taxa reach their highest abundance in the coldest region in the data set. *Derotanypus* is found only once outside this region, at very low abundance (1.5%) in the Lena River Delta which has T July 11.3^{o}C. Although not very abundant *Limnophyes* is found in 63% of the lakes in this region. Endemic for the coldest region are rare taxa *Chaetocladius dentiforceps*-type, *Chaetocladius* type B and *Tvetenia* sp. (not given in the Fig. 3).

The coldest lake in our data set is the lake in Bunge Land, a part of Kotelnyj island, Novosibirsk islands (T July 1.8^{o}C, Laptev, Fig. 1). The chironomid community here is dominated by *Orthocladius* type I and *Hydrobaenus conformis*-type.

The T July interval between 5.0 and 10.0 °C is represented by one lake from Lena-Amga-interfluve from the ES data set, CY (Nazarova et al., 2011) and lakes from three regions previously not included in WS and ES data sets: Kolyma, Indigirka and Kamchatka. In total, 100 taxa are identified from lakes in these regions. Several rare taxa are found only in the lakes of this T July range: *Doncricotopus, Metriocnemus*, *Parachironomus vitosus*-type, *Pseudodiamesa*, and *Tanytarsus nemorosus*-type (not in the Fig. 3).

Orthocladius type I, *P. austriacus*-type and *Derotanypus* are associated with mean July air temperatures of less than 12.0 °C. *Heterotrissocladius grimshawi*-type, *H. maeaeri*-type, *Corynocera oliveri*-type, *Z. zalutschicola*, *Z.* type B, *Limnophyes* and *Micropsectra insignilobus*-type are predominately associated with mean July air temperatures of less than 15.0 °C (Fig. 3). *C. plumosus*-type, *Cricotopus laricomalis*-type, *Einfeldia natchitocheae*-type and *D. nervosus*-type are more abundant in lakes with T July greater than 15.0 °C.

The warmest part of the gradient (≥16.0°C) is represented by 72 lakes from CY, in which we identified 125 taxa. Chironomid communities are dominated by Chironomini. Several rare taxa are restricted mainly to the warmest part of the gradient: *Microchironomus, Einfeldia dissidens*-type, *Glyptotendipes pallens*-type, and *Labrundinia* (Fig. 3). Rare taxa *Xenochironomus, Stenochironomus, Rheotanytarsus*, *Rheocricotopus effusus*-type, *Paratendipes, Paralauterborniella*, *Eukiefferiella devonica*-type, and *Acamptocladius* are entirely restricted to the lakes with T July greater than 16.0 °C (not in the Fig.3).

The combined data set includes 11 geographical areas: Komi, Pechora, Putorana, Anabar, Lena, CY, NY, Laptev, Indigirka, Kolyma, Kamchatka (Fig. 1). DCA analysis and squared chord distance analysis show that the most taxonomically distinct regions are

Laptev, Komi, CY and NY (Fig. 4, Table 2). The regions with the lowest TD are those in the middle of the study area: Pechora, Putorana and Anabar (TD=0.68). Kamchatka has the highest similarity to region Lena (TD=0.88) and to the geographically closest region Kolyma (TD=0.95) and shows the lowest similarity to the NY and CY (TD=1.35 and 1.24).

4.2.3. Comparison of taxonomic composition and taxon responses to July air temperature in the WS and ES data sets

Of 177 taxa in total, 20 taxa are found only in the WS data set and 23 taxa only in the ES data set. Except for *Mesocricotopus, Parachaetocladius, Potthastia* type 2, the rest of the taxa, which are found in only WS data set, had single occurrences at low abundances.

The two data sets have similar dominant taxa but differ in the frequency of occurrences of these taxa. More common in the WS data set than in the ES data set are *Ablabesmyia, Procladius, Stictochironomus, C. mancus*-type, *Stempellinella – Zavrelia, Tanytarsus pallidicornis*-type. In the ES data set *C. anthracinus*-type, *Cladopelma, Parachironomus varus*-type, *Polypedilum nubeculosum*-type, *C. ambigua, Corynoneura arctica*-type, *C. laricomalis*-type, *Limnophyes – Paralimnophyes,* and *Ps. sordidellus*-type are more frequent than in the WS data set.

In the combined data set 86.1 % of the non-rare taxa have a statistically significant response to T July.

The T July optima of each taxon in the WS and ES data sets are similar and all differences in optima, apart from *Alabesmyia*, lie within the taxon-specific T July

tolerance. Although it is within the range of tolerance, the biggest difference in T July optima between the models was the estimate for *P. penicillatus*-type: 17.1°C in WS data set, 11.5°C in ES data set.

5. Ordination of the full set of data

VIF and CCAs show that latitude, longitude, T Jan, altitude and continentality are intercorrelated and are subsequently eliminated from the analysis. A set of CCAs constrained to individual environmental variables and Monte Carlo permutation tests reveal that four variables explain significant proportions ($p<0.05$) of variance in the data set: T July explains 5.1% of the variance in the data, conductivity 3.7%, pH and WD explain 2.8% each. CCA with these four variables had CCA axis 1 of 0.185 and CCA axis 2 of 0.146, explaining 5.2% and 4.1% of variance in the data, respectively (Table 3). Juggins (2013) suggests that a ratio of eigenvalues of CCA axes 1 and 2 ($\lambda1/\lambda2$) below 1 indicates that potential factors affecting assemblages besides the explored variables have not been assessed. In this case, this ratio is 1.27 ($\lambda1/\lambda2 = 0.185/0.146$), which indicates that the most important explanatory variables are most likely included, although the strong secondary axis suggests that additional explanatory variables may not have been included in the CCA.

CCA axis 1 most strongly correlates with T July (ESM 1). CCA axis 2 correlates with conductivity, T July and WD. Axis 3 correlates with WD and axis 4 shows correlation with pH, T July and WD. From the above we conclude that T July is the environmental variable that explains the largest proportion of the variance in the data and that T July can be used to develop inference models.

Fig. 5 (a-b) presents correlation biplots of the CCA ordination based on 268 lakes, 177 taxa and four significant environmental variables. A CCA biplot of the sample scores shows that sites are mostly grouped by geographical location (Fig. 5a). The bottom left group are tundra lakes situated in the coldest part of our data set: Bol'shoy Lyakhovsky and Bunge Islands (region Laptev) and the Kolyma and Indigirka regions. In the top left are lakes from Pechora, Putorana, Anabar and Kamchatka in forest tundra and taiga zones. CY and Komi lakes that constitute the warmest part of the gradient are grouped in the right part of the biplot. Lakes from NY and Lena River basin are in the centre.

The distribution of chironomid taxa along the CCA axes reflects their ecological spectra (Fig. 5b). The CCA results suggest *Mesocricotopus*, *H. maeaeri*-type and *H. grimshawi*-type, *Monodiamesa, Phaenospectra flavipes*-type are associated with cold deep lakes (Fig. 5b). Taxa typical of cold and shallow lakes are located in the bottom left quadrant: *Derotanypus* (the most cold-stenothermic taxon), *Orthocladius* type-I, *Allopsectrocladius, Acricotopus, Trissocladius* and *Metriocnemus eurynotus*-type. The two right quadrants include taxa characteristic of warmer, continental conditions of CY that are subject to a negative water balance with many lakes having high ion concentrations and higher conductivity. It is difficult to distinguish if taxa, such as *Einfeldia, C. plumosus*-type, *Endochironomus albipennis*-type and *C. laricomalis*-type, are responding to mean July air temperatures and/or pH and conductivity as they are most abundant in the CY lakes, the region with the warmest July air temperatures and many lakes with high conductivity.

6. Development of inference models

T July models with all 268 lakes and 174 taxa (the full dataset model, or FM) yielded relatively high coefficients of determination (r^2_{jack} = 0.66–0.73), high root mean squared errors of prediction (RMSEP = 2.3–2.7) and max biases $_{jack}$ (2.28–3.50) (Table 4, Fig. 6a).

In order to improve statistical parameters of the FM we used first an ecological approach to lake selection. For the 193 lakes reduced data set, which we refer to as the North Russia (NR), we found the same significant environmental variables as in the full data set (T July, Cond, WD and pH). T July was the environmental variable explaining the highest proportion (6.0 %) of the variance in the reduced NR data set, conductivity explains 3.4%, WD 3.2% and pH 2.9%. A CCA with these four significant variables has axes 1 and 2 explaining 6.0% and 3.7%, respectively, of the variance in the NR data set (Table 3, Fig. 5c-d) and also has a higher $\lambda 1/\lambda 2$ of 1.65.

CCA axis 1 in the NR model most strongly correlates with T July (ESM 1). Axis 2 correlates with conductivity, T July and WD. Axis 3 correlates with pH and conductivity and axis 4 shows correlation with WD.

The NR T July air temperature transfer function was made, based on 193 lakes and 162 taxa. Performance statistics indicated the reduced WA-PLS 2-component model (Table 4) performed better than the 268-lakes model, has higher coefficient of determination (r^2_{jack} = 0.87) and lower RMSEP (1.35 ^{o}C). The model over-predicts temperatures below 10.0 ^{o}C, under-predicts above 14.0 ^{o}C and has a maximum bias of 2.21^{o}C (Fig. 6b).

The geographical approach for selecting lakes resulted in the Far East (FE) data set that has 88 sites with a temperature range of 1.8 - 13.3 ^{o}C. In the FE data set the same four environmental variables were significant. A set of CCAs constrained to the individual

environmental variables indicated that T July explains the highest proportion (7.9%) of the variance in the FE data set, WD explains 7.0% of the variance, conductivity 3.8%, pH 2.1%.

CCA with these four significant environmental variables indicated that CCA axes 1 and 2 explain 9.1% and 4.3% of the variance in the data, respectively, and have a $\lambda 1/\lambda 2$ ratio of 2.03 (Table 3, Fig. 5e-f). The 2-component WA-PLS FE model has $r^2_{jack} = 0.81$, RMSEP = 1.3 ^{o}C (Table 4 and Fig. 6c). The model over-predicts temperatures below 8 ^{o}C, under-predicts between 8 and 10.5 ^{o}C.

7. Taxon specific T July optima in the NR and FE data sets

In the FM data set 94 taxa have more than 10 occurrences, 89 of them have more than 10 occurrences in the reduced NR data set and 49 in the FE data set. The generalized linear models demonstrate that in the full data set 86.1% of the taxa have a significant relationship to T July (63.8% of the taxa have a highly significant relationship ($p \leq 0.001$) and 22.3% have a significant relationship ($p \leq 0.05$)). 76.5% of the taxa had a significant relationship to T July in NR model and 79.6% of taxa in FE model (Table 5, ESM 3). Mean T July optima ranged from 3.6 ^{o}C (*Derotanypus*) to 17.9 ^{o}C (*E. dissidens*-type). Comparison of the T July optima in all three models show that in most of cases optima are identical or very similar in both FM and NR models, or slightly lower in the NR than in FM. The exceptions are *Nanocladius branchicolus*-type, *Orthocladius* type-S, *Paratanytarsus* and *T. glabrescens*-type which have optima between 0.7 -1.1 ^{o}C higher in the NR model.

In contrast to high similarity of the taxon-specific T July optima in FM and NR data sets, T July optima of the same taxa in the FE data set are considerably lower with differences

reaching 5.6 °C (ESM 3). T July optima for *C. anthracinus*-type, *C. plumosus*-type, *C. ambigua, C. arctica*-type, *D. nervosus*-type *P. nubeculosum*-type, *C. laricomalis*-type, *P. penicillatus*-type, *Gl. pallens*-type in the FE model are 3.3 to 4.8 °C lower than in the NR model. For *Gl. barbipes*-type this difference reaches 5.6 °C (Tol 2.4 °C). Only for *Derotanypus* and *Psectrocladius barbimanus*-type are T optima equal in all three models. Comparison of the T July optima of non-rare taxa (with more than 10 occurrences in the lakes) between the Norwegian, WS (Self et al., 2011), ES (Nazarova et al., 2011) and combined models (FM, NR and FE) has shown that for 38 of 40 taxa the highest values of T July optima are produced by the WS and ES models (ESM 3, Fig. 7). *H. maeaeri*-type has the highest optimum in the FM and *Z.* type B has the highest optimum in the Norwegian model. *O. oliveri*-type, *Parakiefferiella triquetra*-type and *Z.zalutschicola*-type have small variations of the optima between data sets.

The lowest optima for 36 of 40 taxa are generated by the FE models. But *C. oliveri*-type, *Micropsectra radialis*-type, *C. laricomalis*-type and *O. oliveri*-type, have slightly lower T July optima in the Norwegian model. For *P. penicillatus*-type a very high optimum (17.1°C) was generated only by WS model, whereas Norwegian, ES, FM and NR models produced comparable values (11.2 – 12.4°C). Noticeably lower are the optima generated by both Norwegian and FE models in comparison to more continental WE, ES, FM and NR models for *C. ambigua, C. oliveri*-type, *Stempellinella –Zavrelia, T. lugens*-type, *C. laricomalis*-type, *H. maeaeri*-type, *H. conformis*-type. For *H. conformis*-type Norwegian and FE models gave nearly equal values (5.2 and 5.4 °C), whereas the other models produced much higher optima varying between 10.4 to11.8°C.

8. Discussion

8.1. Taxonomic composition of the data sets

In this study we combined data from geographically remote areas. The distance between the most western (Komi, lake K7) and the most eastern (Kolyma, lake KO29) sampling sites is more than 5290 km. Two data sets and 55 new sites were merged based on several preconditions: high taxonomic similarity of the combined data sets; distribution of chironomid taxa in both data sets is driven mainly by the same ecological factors; responses of individual taxa to the measured environmental variables are similar. High taxonomic similarity was found between the WS and ES data sets (TSI 0.89). In the combined data set the most taxonomically distinctive regions are the coldest region (Laptev), the most western region (Komi) and the continental areas of NY and CY. The highest taxonomic similarity is between the east-west neighbouring Pechora, Putorana and Anabar, and between Kamchatka and Kolyma. Sites with the most oceanic climate and lowest CI have high taxonomic similarity (Kamchatka and Lena) and the low similarity with the most continental regions NY and CY.

There are few examples of faunistic studies of chironomids in the northern arctic and sub-arctic Russia (Makarchenko, 1998; Makarchenko, Makarchenko, 2001; Krasheninnikov, 2013; Krashenninnikov and Gavrilo, 2013). Our study provides new information on the taxonomic diversity of chironomid fauna, biogeographical distribution and can supplement existing information on chironomid biodiversity and ecology in northern Russia. *Cladopelma, C. anthracinus*-type, *C. mancus*-type, *C. ambigua, C. arctica*-type, *P. penicillatus*-type, *Ps. sordidellus*-type, *Procladius*, *T. pallidicornis*-type and *T. lugens*-type, appear to be widespread throughout the whole northern Russia. The taxa may either have a cosmopolitan distribution due to wide environmental tolerances, or the

morphotype may comprise different species in different regions and, therefore, possibly different temperature optima between geographical regions.

Widespread in the data set are *Limnophyes* – *Paralimnophyes* and *Z. zalutschicola*. *Zalutschia* and some species of *Limnophyes* are associated with aquatic macrophytes and *Limnophyes* typically occurs in very shallow water (Kansanen, 1985; Hofmann, 1998). The majority of Russian lakes are from gently undulating tundra, so thawing of permafrost may result in increased water depth and lake area, thereby increasing the size of the littoral zone and the area available for macrophyte growth and therefore may account for the wide distribution of these taxa.

Orthocladiinae are known to be dominant in all Russian Arctic Islands (Novaya Zemlya, Franz Josef Land, Severnaya Zemlya, Novosibirsk Islands) (Krasheninnikov, 2013; present study). In modern times 39 species are recorded in Novaya Zemlya (T July 4.6^{o}C), including *P. austriacus* (Kieffer, 1924) (Makarchenko, 1985; Makarchenko et al., 1998), and36 species are found in Wrangel Island (T July 2.7^{o}C) (Makarchenko, 1998; Makarchenko, Makarchenko, 2001). Krashenninnikov and Gavrilo (2013) have recorded 4 species in the extreme environments of Franz Josef Land (Mean T July +0.7^{o}C): *Chaetocladius* (*Amblycladius*) *franzjosephiensis* Krasheninnikov, *Metriocnemus* (*Metriocnemus*) *eurynotus* (Holmgren, 1883) and *M.* (*M.*) *sibiricus* (Lundström, 1915). Holarctic species are widespread in all islands, apart from Novosibirsk islands, where only 8 of 17 known species have Holarctic distribution. To date 17 chironomid species were known from Novosibirsk Islands (Krashenninnikov, 2013), among which *Acricotopus lucens* (Zetterstedt, 1850), *Heterotrissocladius subpilosus* (Kieffer, 1911), *Derotanypus limbatus* (Lundström, 1915), *Procladius* (*Holotanypus*) *culiciformis*

(Linnaeus, 1767), *Orthocladius* (*Pogonocladius*) *consobrinus* (Holmgren, 1869) can be attributed to morphotypes that are found in our study. We recorded 51 taxa in the Novosibirsk Islands. Although taxonomic resolution of palaeoecological study is coarser than modern faunistic investigations, our findings demonstrate that taxonomic diversity in such harsh environments as Novosibirsk Islands is higher than was described for the Islands before.

Earlier surveys have demonstrated some similarities between the chironomid fauna of Far East and other parts of Siberia (Karationis et al., 1956; Ogay, 1979; Salova, 1993; Kiknadze et al., 1996). We observed a relatively high abundance of head capsules from the subfamily Diamesinae in Kamchatka, whereas almost none were found in other regions included in the data set. Makarchenko et al. (2005) recorded 52 Diamesinae taxa in a catalogue of chironomid fauna of the Russian Far East. High taxonomic richness and abundance of Diamesinae may be a distinctive feature of the region. In total previous investigations (Makarchenko et al., 1999, 2005; Zorina, 2001, 2003;2006a,b; 2013) recorded 74 species for Kamchatka, which is comparable with the taxonomic richness that was found in our investigation. Most recorded species (60%) are Palaearctic and 40% of all species have Holarctic distribution.

8.2. Environmental parameters

In both combined data sets similar environmental parameters played an important role in the distribution of chironomids. In the WS training set chironomid assemblages are significantly correlated with T July, CI, WD and precipitation. In the ES data set the significant environmental variables were T July, WD, conductivity, Secchi depth, Ca^{2+}

pH and concentrations of Mn, Ca, Cl and Fe. The ordinations suggest that there is a strong relationship between the distribution and abundance of chironomids and T July, conductivity, pH and WD in the combined data set. These strong relationships also appear to persist over the wider geographical area of north-west Eurasia (i.e. Norway and north-west Russia), are well documented and have been used to develop chironomid-inferred temperature transfer functions in northern Eurasia (Olander et al., 1999; Brooks and Birks, 2001; Larocque et al., 2001; Luoto, 2008) as well as in northern North America (Palmer et al., 2002; Walker and Cwynar, 2006).

Unfortunately, measurements of nutrients and dissolved oxygen concentrations are only available for a limited number of lakes in our combined data set, which does not allow us to assess the relative importance of these variables. Previous studies have shown that nutrients and oxygen can have an important effect on chironomid assemblages (e.g. Lotter et al., 1998; Quinlan and Smol, 2001b; Langdon et al., 2006). Nutrients and oxygen are often of secondary importance to chironomid assemblages when large temperature gradients are considered, which can be partly explained by the correlation between temperature and nutrients (Eggermont and Heiri, 2012). The large temperature gradient in our combined data and the fact that T July was the environmental variable that explained most variance in the data, we assume that T July is the most important environmental variable for the chironomid distributions in the investigated lakes.

The combined FM data set includes a greater number of lakes and chironomid taxa than in any other published data set (Nazarova et al., 2011, ESM 3) and crosses the longest sampled T July gradient. But models developed from chironomid training sets benefit not only from the number of sampled sites and the length of the gradient, but also from the

evenness of the distribution of sampled lakes along that gradient. This has remained a problem in the WS and especially in the ES data set. Uneven sampling of the environmental gradient can have a negative influence on transfer-function performance. Ter Braak and Looman (1986) demonstrated that the efficiency of weighted averaging (WA) for estimating species' optima and tolerances is high only when the environmental gradient is evenly sampled. Poorly estimated WA optima are unlikely to give the most reliable reconstructions, so that training sets with evenly sampled gradients should perform better than those with unevenly sampled gradients (Telford and Birks, 2011). Although Ginn et al. (2007) found that the cross-validation performance statistics from an uneven data set and a uniform data set were similar, suggesting that distribution of the lakes along the gradient of interest is not vital, the evenness of distribution has ecological and statistical sense, because species optima are more precisely known (and more analogues are available) in the part of the gradient with most observations. Hence optima estimates are most precise here and compensate for the less precise estimates in the less well sampled parts of the gradient (Telford and Birks, 2011).

Distribution of the sampled lakes along the temperature gradient was significantly improved by merging of the data sets and by inclusion of lakes from regions not previously included in either dataset filling the gaps in the T July gradient.

8.3. Taxa - T July relationship

The share of the non-rare taxa positively responding to T July is similar in both WS and ES data sets and in the combined data set (85, 87 and 86%, respectively). T July optima in both data sets are similar and differences mostly lay within the calculated

taxon-specific T July tolerances. The biggest difference in T July optima was estimated for *P. penicillatus*-type (WS/ES = 17.1/11.5 ^{o}C) and can be attributed to the broader distribution of the taxon in the ES data set, where a longer T July gradient, including the coldest region Laptev, was sampled. *Paratanytarsus penicillatus*-type shows a highly significant response to T July in the ES data set, whilst in the WS data set, *P. penicillatus*-type has HOF response model type II (a sigmoidal increasing response), which suggests that the coldest part of T July range of the taxon was not sampled in the WS data set to give a unimodal response for this taxon. This is supported by earlier biogeographical studies that show that *P. penicillatus*-type is widely distributed in cold environments of northern Russia: both *P. penicillatus*, and *P. austriacus*, were found in seven Hibin lakes in the Kola Peninsula (T July 13.2^{o}C) (Zelentsov, 2009), in Kamchatka (8.0-11.0 ^{o}C) (Makarchenko et al., 2005) and in Tiksi, Lena Delta (T July 6.8^{o}C) (Zelentsov, Shilova, 1996).

Taxon specific T July optima are very similar or identical in the FM and NR models, but the FE model produced considerably lower optima, with the differences reaching 5.6 ^{o}C. The low T July optima generated by the FE model may be underestimated because of the absence of warm areas within the region. This may bias the reconstruction. The FE model was applied for palaeoclimate reconstructions from sediments of several lakes in Kamchatka (Solovieva et al.; Self et al, this volume).

Continentality was excluded from the analysis due to its strong correlation to T July. However for 95% of the taxa the highest values of T July optima are generated by the more continental WS and ES models and the lowest optima for 90% of the taxa are generated by the Norwegian and FE models that include more sites with low

continentality. 50 % of the taxa have, similar T July optima (≤ 1°C) in the Norwegian and reduced NR model. Probably this reduction of T July optima in the new NR model is caused by a more even distribution of the lakes along the temperature gradient. The higher optima for the WS and ES model may reflect the gradient length and distribution of lakes along the July air temperature gradient in both data sets and the influence of continentality.

8.4. T July inference models

FM T July WA-PLS models with all 268 lakes yielded reasonably high coefficients of determination (Table 4) but very high RMSEP when compared with other chironomid-based mean July air temperature inference models (Nazarova et al., 2011, ESM 3). The FM has RMSEP = 2.3 °C that constitutes 15 % of the sampled gradient, which is similar or even lower than in some chironomid-based air temperature inference models (Larocque et al., 2001; Palmer et al., 2002; Rosenberg et al., 2004; Woodward and Shulmeister, 2006; Dieffenbacher-Krall et al., 2007). However implementation of the model with high RMSEP brings additional uncertainty into the reconstructions and may dwarf any reconstructed temperature fluctuations.

In order to improve statistical parameters of the 268 lakes FM data set and to produce an applicable model for T July reconstructions in northern Russia we used ecological and geographical approaches for lake selection. An ecological approach implies exclusion from the analysis of lakes that increase variability of parameters other than the T July (Brooks and Birks, 2000). Following this approach we excluded from the analysis lakes

with high conductivity and pH. Deleting the samples with low pH did not influence the statistical parameters of the model.

A geographical approach implies selection of the lakes based on their relative geographical proximity. In our case we selected a data set of lakes which could produce a T July model potentially better applicable in the Russian Far East. For this purpose we chose a natural geographical break in the sampled data set along the Verkhoyansk Mountains in north-eastern Siberia (Rakovskaya and Davydova, 2001). This geographical approach is based on the thesis that inference models have limited application outside of the regions in which they were developed because differences in faunal composition between sites in the calibration set and the fossil sequence make data difficult to interpret and results are sometimes unreliable (Lotter et al., 1999; Andreev et al., 2004; Andreev et al., 2005). However, the use of transfer functions from geographically distinct areas may be justified when fossil taxa are poorly represented in local training sets (Larocque-Tobler, 2010).

Using the both approaches we developed NR and FE training sets and statistical models for T July reconstructions. In both models T July is the main ecological parameter driving the abundance and distribution of chironomids, both models have a high $\lambda 1/\lambda 2$ ratio, r^2_{jack} (0.87 and 0. 81), and relatively low RMSEP (1.35 and 1.3 ^{o}C). In the NR model 76.5% of the taxa had a significant relationship to T July and in FE model 79.6% of taxa had a significant relationship to T July.

Given the fact that all regional parts of the data set have high taxonomic similarity and the models have good statistical parameters, we suppose that NR and FE models can be

applied for palaeoclimatic reconstructions in most parts of the northern and north-eastern Russia.

9. Conclusions

- Following taxonomic standardisation between chironomid WS and ES calibration sets we have merged the data sets and added 55 lakes from three new East Siberian regions to the combined data set.
- High taxonomic similarity was found between the WS and ES data sets. The most taxonomically distinctive regions in the combined data set are the coldest region Laptev, and the continental areas of CY, NY and Komi. The highest taxonomic similarity was found between the neighbouring regions and between the sites with the most oceanic climate. More continental areas differ considerably from the sites with oceanic climate.
- The analysis of chironomid taxa, and environmental data sets from 268 lakes has shown that T July explained the highest variation in the chironomid species data.
- Full set T July Model (FM) with all 268 lakes yielded high coefficients of determination (r^2_{jack} = 0.67–0.73), high root mean squared errors of prediction (RMSEP = 2.3–2.8) and max biases $_{jack}$ (2.28–3.48).
- In order to improve statistical parameters of the FM we used two approaches: an ecological and a geographical approach to lake selection. Based on ecological approach, we deleted all the lakes with conductivity above 500 µS/cm and with pH ≥8. The NR T July air temperature transfer function was constructed, based on 193

lakes. Reduced WA-PLS 2-component NR model has higher coefficient of determination (r_{jack} 2 = 0.87) and lower RMSEP (1.35 ^{o}C).

- To produce a T July model that can have a potentially better applicability in the Russian Far East (FE) we used geographical approach. For this we ranked all the lakes by longitude and included into the FE training set only sites east of longitude 140 ^{o}E, corresponding to the Verkhoyansk Mountains. The FE training set includes 88 sites with a T July range of 1.8 - 13.3 ^{o}C. The 2-component WA-PLS FE model has the r^2_{jack} of 0.81, RMSEP = 1.3 ^{o}C.
- Comparison of the T July optima of taxa with more than 10 occurrences in the lakes between the Norwegian, WS, ES and combined models (FM, NR and FE) has shown that for 38 of 40 taxa the highest values of T July optima are produced by the WS and ES models. Noticeably lower for the most of the taxa are the optima generated by both Norwegian and FE models in comparison to more continental WE, ES FM and NR models, which may reflect the gradient length and distribution of lakes along the July air temperature gradient and the influence of continentality.

10. Acknowledgements

We thank all Russian and German colleagues who helped us during the fieldwork in north-eastern Russia over several years. Many thanks to Antje Eulenberg and Ute Bastian (AWI) for the help in laboratory work, to Dr. Evgenia Vinogradova for the help with chironomid slides preparation. Sincere thanks to the reviewers for their valuable comments. This study was supported by Alexander von Humboldt Foundation, Deutsche Forschungsgemeinschaft (DFG), and by the Subsidy allocated to Kazan Federal University for the state assignment in the sphere of scientific activities

References

Altman, D., Bland, J.M., 1983. Measurement in medicine: the analysis of method comparison studies. The Statistician, 32, 307–317.

Andreev, A., Tarasov, P., Schwamborn, G., Iljashuk, B., Iljashuk, E., Bobrov, A., Klimanov, V., Rachold, V., 2004. Holocene palaeoenvironmental records from Nikolay Lake, Lena Delta, Arctic Russia. Palaeogeography Palaeoclimatology Palaeoecology, 209, 197-217.

Andreev, A., Tarasov, P.E., Ilyashuk, B.P., Ilyashuk, E.A., Cremer, H., Hermichen, W.D., Hubberten, H.-W., Wischer, F., 2005. Holocene environmental history recorded in the Lake Lyadhej-To sediments, Polar Urals, Russia. Palaeogeography Palaeoclimatology Palaeoecology, 223, 181-203.

Barley, E.M., Walker, I.R., Kurek, J., Cwynar, L.C., Mathewes, R.W., Gajewski, K., Finney, B.P., 2006. A northwest North American training set: distribution of freshwater midges in relation to air temperature and lake depth. Journal of Paleolimnology, 36, 295-314.

Behar, S., 1997. Testing the Waters: Chemical and Physical Vital Signs of a River. Montpelier, Vermont: River Watch Network.

Birks, H.J.B., 1995. Quantitative palaeoenvironmental reconstructions. In: Maddy D, Brew JS (Eds) Statistical modelling of quaternary science data. Technical Guide 5, Quaternary Research Association, Cambridge, pp. 161-254.

Birks, H.J.B., 1998. Numerical tools in paleolimnology progress, potential, and problems. Journal of Paleolimnology, 20, 307-332.

Birks, H.J.B., Juggins, S., Line, J.M., 1990a. Lake surface-water chemistry reconstructions from palaeolimnolical data. In: Mason B (Ed.) The surface water acidification programme. Cambridge University Press, Cambridge, pp 301–313.

Birks, H.J.B., Line, .J.M., Juggins, S., Stevenson, A.C., ter Braak, C.J.F., 1990b. Diatoms and pH reconstruction. Philosophical Transactions of the Royal Society B: Biological Sciences, 327, 263-278.

Brooks, S.J., 2006. Fossil midges (Diptera: Chironomidae) as palaeoclimatic indicators for the Eurasian region. Quaternary Science Reviews, 25, 1894-1910.

Brooks, S.J., Birks, H.J.B., 2000. Chironomid-inferred late-glacial and early-Holocene mean July air temperatures for Kråkenes Lake, western Norway. Journal of Paleolimnology, 23, 77-89.

Brooks, S.J., Birks, H.J.B., 2001. Chironomid-inferred air temperatures from late-glacial and Holocene sites in north-west Europe: progress and problems. Quaternary Science Reviews, 20, 1723- 1741.

Brooks, S.J., Langdon, P.G., Heiri, O., 2007. Using and identifying chironomid larvae in palaeoecology. QRA Technical Guide № 10, Quaternary Research Association, London.

Coffman, W., Ferrington L., 1996. Chironomidae. In: Merritt R., Cummins K. (Eds.) An Introduction to the Aquatic Insects of North America. Dubuque, Iowa, USA: Kendall/Hunt Publishing Company, pp. 591-754.

Dieffenbacher-Krall, A.C., Vandergoes, M.J., Denton, G.H., 2007. An inference model for mean summer air temperatures in the Southern Alps, New Zealand using subfossil chironomids. Quaternary Science Reviews, 26, 2487-2504.

Dowsett, H.J., Robinson, M.M., 1998. Application of the modern analogue technique (MAT) of sea surface temperature estimation to middle Pliocene North Pacific planktic foraminifer assemblages. Palaeontologia Electronica, 1, 1.

Duff, K., Laing, T.E., Smol, J.P., Lean, D.R.S., 1998. Limnological characteristics of lakes located across arctic tree-line in northern Russia. Hydrobiologia, 391, 205-222.

Eggermont, H., Russel, J.M., Schettler, G., Van Dame, K., Bessems, I., Verschuren, D., 2007. Physical and chemical limnology of alpine lakes and pools in the Rwenzori Mountains (Uganda-DR Congo). Hydrobiologia, 592, 151-173.

Eggermont, H., Heiri, O., 2012. The chironomid-temperature relationship: expression in nature and palaeoenvironmental implications. Biological Reviews, 87, 430-456.

Engels, S., Self, A.E., Luoto, T.P., Brooks, S.J., Helmens, K.F., 2014. A comparison of three Eurasian chironomid–climate calibration datasets on a W–E continentality gradient and the implications for quantitative temperature reconstructions. Journal of Paleolimnology, 51, 529-547.

Gavrilova, M.K., 1998. Klimaty kholodnyh regionov zemli (Climates of cold regions of the world) Russian Academy of Science. Russian Academy of Science, Siberian Branch, Permafrost Institute, Yakutsk. [In Russian]

Gavin D.G., Oswald W.W., Wahl E.R., Williams J.W., 2003. A statistical approach to evaluating distance metrics and analog assignments for pollen records. Quaternary Research, 60, 356–367.

Ginn, B.K., Cumming, B.F., Smol, J.P., 2007. Diatom-based environmental inferences and model comparisons from northeastern North American lakes. Journal of Phycology, 43, 647–661.

Gorczynski, W., 1920. Sur le calcul du degre du continentalisme et son application dans la climatologie. Geografiska Annaler, 2, 324-331.

Grieser, J., Gommes, R., Cofield, S., Bernardi, M., 2006. Data Sources for FAO Worldmaps of Koeppen Climatologies and Climatic Net Primary Production. The Agromet Group, SDRN, Food and Agriculture Organization of the United Nations, Rome, Italy.

Hammer, Ø., Harper, D.A.T., Raan, P.D., 2001. PAST: Palaeontological statistics software package for education and data analysis. Palaeontologia electronica, 41, 9 p.

Hofmann, W., 1998. Cladocerans and chironomids as indicators of lake level changes in north temperate lakes. Journal of Paleolimnology, 19, 55-62.

IFM-GEOMAR, 2008. BMBF-Verbundvorhaben KALMAR: Kurile-Kamchatka and Aleutean Marginal Sea -Island Arc Systems: Geodynamic and Climate Interaction in Space and Time; Zwischenbericht 2006/2007, IFM-GEOMAR, Kiel.

Ilyashuk, E., Ilyashuk, B., Hammarlund, D., Larocque, I., 2005. Holocene climatic and environmental changes inferred from midge (Diptera: Chironomidae, Chaoboridae,

Ceratopogonidae) records at Lake Berkut, southern Kola Peninsula,Russia. The Holocene, 15, 897-914.

Jones, V.J., Solovieva, N., Self, A.E., McGowan, S., Rosen, P., Salonen, J.S., Seppa, H., Valiranta, M., Parrott, E., Brooks, S.J., 2011. The influence of Holocene tree-line advance and retreat on an arctic lake ecosystem: a multi-proxy study from Kharinei Lake, North Eastern European Russia. Journal of Paleolimnology, 46, 123–137.

Juggins, S., 2007. C2 Version 1.5 User guide. Software for ecological and palaeoecological data analysis and visualisation. Newcastle University, Newcastle upon Tyne, UK.

Juggins, S., 2013. Quantitative reconstructions in palaeolimnology: new paradigm or sick science? Quaternary Science Reviews, 64, 20-32.

Kansanen, P.H., 1985. Assessment of pollution history from recent sediments in Lake Vanajavesi, southern Finland. II. Changes in Chironomidae, Choaboridae and Ceratopogonidae (Diptera) fauna. Annales Zoologici Fennici, 22, 57-90.

Kienast, F., Wetterich, S., Kuzmina, S., Schirrmeister, L., Andreev, A., Tarasov, P., Nazarova, L., Kossler, A., Frolova, L., Kunitsky, V., 2011. Paleontological records prove boreal woodland under dry inland climate at today's Arctic coast in Beringia during the last interglacial. Quaternary Science Reviews, 30, 2134-2159.

Kiknadze, I.I., Istomina, A.G., Gunderina, L.I., Salova, T.A., Ajmanova, K.G., Savvinov, D.D., 1996. Kariofondy chironomid kriolitozony Yakutii: triba Chironomini (Cariofonds of chironomids of Criolithozone of Yakutia: triba Chironomini). Nauka, Novosibirsk. [In Russian]

Krasheninnikov, A.B. 2013. Predvaritelnye dannye po faune i rasprostraneniju komarov-zvonzov (Diptera, Chironomidae) Rossijskogo sektora arktiki. (Preliminary data on fauna and distribution of chironomids in Russian arctic). Vestnik Permskogo Universiteta, 1, 32-36. [In Russian]

Krasheninnikov, A.B., Gavrilo, M.V., 2013. New data on the chironomids (Diptera, Chironomidae, Orthocladiinae) of Franz Joseph Land Archipelago. Euroasian Entomological Journal, 12, 157–160. [In Russian]

Krestov, P.V., Omelko, A.M., Nakamura, Y., 2008. Vegetation and natural habitats of Kamchatka. Berichte der Reinhold-Tüxen-Gesellschaft, Hannover, 20, 195–218.

Kumke, T., Ksenofontova, M., Pestryakova, L., Nazarova, L., Hubberten, H.W., 2007. Limnological characteristics of lakes in the lowlands of Central Yakutia, Russia. Journal of Limnology, 66, 40-53.

Langdon, P.G., Ruiz, Z.O.E., Brodersen, K.P. Foster, I.D.L., 2006. Assessing lake eutrophication using chironomids: understanding the nature of community response in different lake types. Freshwater Biology, 51, 562-577.

Larocque, I. 2001. How many chironomid head capsules is enough? A statistical approach to determine sample size for paleoclimatic reconstruction. Palaeogeography Palaeoclimatology Palaeoecology, 172, 133–142.

Larocque, I., Hall, R.I., Grahn, E., 2001. Chironomids as indicators of climate change: a 100-lake training set from a subarctic region of northern Sweden (Lapland). Journal of Paleolimnology, 26, 307-322.

Larocque-Tobler, I., 2010. Reconstructing temperature at Engelsee, Switzerland using North American and Swedish chironomid transfer functions: potential and pitfalls. Journal of Paleolimnology, 44, 243-251.

Lindquist, S.J., 1999. The Timan-Pechora Basin Province of Northwest Arctic Russia: Domanik - Paleozoic Petroleum System: U.S. Geological Survey Open-File Report 99-50-G. 24 p.

Lotter, A.F., Birks, H.J.B., Hofmann, W., Marchetto, A., 1998. Modern diatom, cladocera, chironomid, and chrysophyte cyst assemblages as quantitative indicators for the reconstruction of past environmental conditions in the Alps. II. Nutrients. Journal of Paleolimnology, 19, 443-463.

Lotter, A.F., Walker, I.R., Brooks, S.J., Hofmann, W., 1999. An intercontinental comparison of chironomid palaeotemperature inference models: Europe vs. North America. Quaternary Science Reviews, 18, 717-735.

Luoto, T.P., 2008. Subfossil Chironomidae (Insecta: Diptera) along a latitudinal gradient in Finland: development of a new temperature inference model. Journal of Quaternary Science, 24, 150-158.

Mackay, A.W., Bezrukova, E.V., Leng, M.J., Meaney, M., Nunes, A., Piotrowska, N., Self, A., Shchetnikov, A., Shilland, E., Tarasov, P., Wang, L., White, D., 2012. Aquatic ecosystem responses to Holocene climate change and biome development in boreal central Asia. Quaternary Science Reviews, 41, 119-131.

Makarchenko, E.A., 1985. Chironomids of the Far East of the USSR. Subfamilies Podonominae, Diamesinae and Prodiamesinae (Diptera, Chironomidae). Vladivostok: FESC AS USSR, p. 1–208. [In Russian]

Makarchenko, E.A., 1998. Chironomids of the subfamily Diamesinae (Diptera, Chironomidae) of North hemisphere (systematics, biology, biogeography): Habilitation dissertation. Vladivostok State University, Vladivostok. 493 p. [In Russian]

Makarchenko, E.A., Makarchenko, M.A., 1999. Chironomidae. In: Tsalolikhin SJ (ed) Key to freshwater invertebrates of Russia and adjacent lands. V.4. Zoological Institute RAS St. Petersburg, pp 210-295, 670-857. [In Russian]

Makarchenko, E.A., Makarchenko, M.A., 2001. Fauna of chironomids of the subfamily Orthocladiinae (Diptera, Chironomidae) from Wrangel Island. Vladimir Y. Levanidov's Biennial Memorial Meetings, 1, 174–186. [In Russian]

Makarchenko, E.A., Makarchenko, M.A., Vehov, N.V., 1998. Preliminary data on chironomid fauna (Diptera, Chironomidae) of the archipelago Novaya Zemla. Novaya Zemla. Nature, history, archaeology, culture. Moscow. p. 262–267. [In Russian]

Makarchenko, E.A., Makarchenko M.A., Zorina O.V., 1999. A preliminary list of Chironomidae (Diptera) of the Primorye territory (Russian Far East). Far Eastern Entomologist, 78, 1-15. [In Russian]

Makarchenko, E.A., Makarchenko, M.A., Zorina, O.V., Sergeeva, I.V., 2005. Preliminary data on fauna and taxonomy of Chironomids (Diptera, Chironomidae) of Russian

Far East. Vladimir Y. Levanidov's Biennial Memorial Meetings, 3, 394–420. [In Russian]

Mukhin, N.I., Petrakova, V.M., Shevchenko, A.I., 1964. Klimaticheskie kxarakteristiki Pechorskogo ugol'nogo basseina. (Climatic characteristics of the Pechora basin). In: Geokrinologicheskie usloviya Pechorskogo ugol'nogo basseina. (Geocrinological conditions of the Pechora coal basin), Moscow, pp. 20–28 [in Russian].

Nazarova, L., 2012. Chironomids in Quaternary permafrost deposits in the Siberian Arctic. Chironomus. Newsletter of Chironomid Research, 25, 39-42.

Nazarova, L., Kumke, T., Pestrjakova, L., Hubberten, H.-W., 2005. Chironomid fauna of Central Yakutian lakes (Northern Russia) in palaeoenvironmental investigation. Chironomus. Newsletter of Chironomid Research, 18: 25-27.

Nazarova, L.B., Pestryakova, L.A., Ushnitskaya, L., Hubberten, H.W., 2008. Chironomids (Diptera: Chironomidae) in lakes of central Yakutia and their indicative potential for paleoclimatic research. Contemporary Problems of Ecology, 1, 335-345.

Nazarova, L., Herzschuh, U., Wetterich, S., Kumke, T., Pestjakova, L., 2011. Chironomid-based inference models for estimating mean July air temperature and water depth from lakes in Yakutia, northeastern Russia. Journal of Palaeolimnology, 45, 57–71.

Nazarova, L., Lüpfert, H., Subetto, D., Pestryakova, L., Diekmann, B., 2013a. Holocene climate conditions in Central Yakutia (North-Eastern Siberia) inferred from sediment

composition and fossil chironomids of Lake Temje. Quaternary International, 290-291, 264-274.

Nazarova, L., de Hoog, V., Hoff, U., Diekmann, B., 2013b. Late Holocene climate and environmental changes in Kamchatka inferred from subfossil chironomid record. Quaternary Science Reviews, 67, 81-92.

Neshataeva, V., 2008. The Plant Cover of the Kamchatka Peninsula (North of the Russian Far East)', Komarov Botanical Institute Russian Press. St.-Petersburg, Russia.

New, M., Lister, D., Hulme, M., Makin, I., 2002. A high-resolution data set of surface climate over global land areas. Climate Research, 21, 1–25.

Ogay, R.I., 1979. Zoobenthos. In: Biologia Viluiskogo vodohranilisha (Biology of Viluy water reservoir). Nauka, Novosibirsk. p. 136-155. [In Russian]

Oksanen, J., Minchin, P.R., 2002. Continuum theory revisited: what shape are species responses along ecological gradients? Ecological Modelling, 157, 119-129.

Olander, H., Birks, H.J.B., Korhola, A., Blom, T., 1999. An expanded calibration model for inferring lakewater and air temperatures from fossil chironomid assemblages in northern Fennoscandia. The Holocene, 9, 279-294.

Oliver, D.R., 1971. Life history of the Chironomidae. Annual Review of Entomology, 16, 211–238.

Overland, J.E., Wang, M., 2005. The Arctic climate paradox: the recent decrease of the Arctic Oscillation. Geophysical Research Letters, 32, 1-5.

Palmer, S., Walker, I.R., Heinrichs, M., Hebda, R., Scudder, G., 2002. Postglacial midge community change and Holocene palaeotemperature reconstructions near treeline, southern British Columbia (Canada). Journal of Paleolimnology, 28, 469-490.

Porinchu, D.F., Cwynar, L.C., 2000. The distribution of freshwater Chironomidae (Insecta: Diptera) across the treeline near the lower Lena River, northeast Siberia. Arctic, Antarctic and Alpine Research, 32, 429-437.

Quinlan, R., Smol, J.P., 2001a. Setting minimum head capsule abundance and taxa deletion criteria in chironomid-based inference models. Journal of Paleolimnology, 26, 327-342.

Quinlan, R., Smol, J.P., 2001b. Chironomid-based inference models for estimating end-of-summer hypolimnetic oxygen from south-central Ontario shield lakes. Freshwater Biology, 46, 1529-1551.

Rakovskaya, E.M., Davydova, M.I., 2001. Physical geography of Russia. Part. 2. Moscow, 304 p.

Rees, A.B.H., Cwynar, L.C., Cranston, P., 2008. Midges (Chironomidae, Ceratopogonidae, Choaboridae) as a temperature proxy: a training set from Tasmania, Australia. Journal of Paleolimnology, 40, 1159-1178.

Renberg, I., 1991. The HON-Kajak sediment corer. Journal of Paleolimnology, 6, 167-170.

Rosenberg S.M., Walker I.R., Mathews R.W., Hallet D.J., 2004. Midge-inferred Holocene climate history of two subalpine lakes in southern British Columbia. Holocene, 14, 258-271.

Salova, T.A., 1993. Zoobenthos. In: Ekologia Srednei Amgi (Ecology of the Middle Amga River). Yakutsk. p. 44-48. [In Russian]

Sarmaja-Korjonen, K., Kultti, S., Valiranta, M., Solovieva, N., 2003. Mid-Holocene palaeoclimatic and palaeohydrological conditions in northeastern European Russia: a multi-proxy study of Lake Vankavad. Journal of Paleolimnology, 30, 415-426.

Self, A.E, Andren, E., Klimaschewski, A., Solovieva, N., Jones, V.J.J., Hammarlund, D., Brooks, S.J. 2014. Olive-backed Lake, Central Kamchatka. Evidence of Holocene environmental change from a mountain lake in central Kamchatka. This volume.

Self, A.E., Brooks, S.J., Birks, H.J.B., Nazarova, L., Porinchu, D., Odland, A., Yang, H., Jones, V.J., 2011. The distribution and abundance of chironomids in high-latitude Eurasian lakes with respect to temperature and continentality: development and application of new chironomid-based climate-inference models in northern Russia. Quaternary Science Reviews, 30, 1122-1141.

Sokal, R.R., Rohlf, F.J., 1995. Biometry: the principles and practice of statistics in biological research. W. H. Freeman and Co: New York.

Solovieva, N., Jones, V.J., Appleby, P.G., Kondratenok, B.M., 2002. Extent, environmental impact and long-term trends in atmospheric contamination in the Usa basin of east-European Russian Arctic. Water, Air and Soil Pollution, 139, 237-260.

Solovieva, N., Jones, V.J., Nazarova, L., Brooks, S.J., Birks, H.J.B., Grytnes, J.-A., Appleby, P.G., Kauppila, T., Kondratenok, B., Renberg, I., Ponomarev, V., 2005.

Palaeolimnological evidence for recent climate change in lakes from the Northern Urals, Russia. Journal of Paleolimnology, 33, 463-482.

Solovieva, N., Jones, V.J., Birks, H.J.B.,Appleby, P.G., Nazarova, L., 2008. Diatom responses to 20th century climate warming in lakes from the Northern Urals, Russia. Palaeogeography Palaeoclimatology Palaeoecology, 259, 96 –106.

Solovieva, N., Jones, VJ., Self, A., Brooks S., Klimaschewski, A.,. Bennett, K, Nazarova, L., Holocene history of Lifebuoy Lake, NW of the Kamchatka Peninsula. This volume.

Sørensen, T. 1957. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. Kongelige Danske Videnskabernes Selskab, 5, 1–34.

Telford, R. J., Birks, H. J. B., 2011. Effect of uneven sampling along an environmental gradient on transfer-function performance. Journal of Paleolimnology, 46, 99–106.

ter Braak, C.J.F., 1990. Update notes: CANOCO Version 3.10. Agricultural Mathematics Group, Wageningen.

ter Braak, C.J.F., 1995. Ordination. In: Jongman RHG, ter Braak CJF, van Tongeren OFR (Eds.) Data analysis in community and landscape ecology. Cambridge University Press, Cambridge. pp. 69–173.

ter Braak, C.J.F., Juggins, S., 1993. Weighted averaging partial least squares regression (WA-PLS): an improved method for reconstructing environmental variables from species assemblages. Hydrobiologia, 269-270, 485-502.

ter Braak, C.J.F., Looman, C.W.N., 1986. Weighted averaging, logistic regression and the Gaussian response model. Vegetatio, 65, 3-11.

ter Braak, C.J.F., Šmilauer, P., 1998. CANOCO reference manual and user's guide to Canoco for Windows (version 4). Centre for Biometry Wageningen, New York.

ter Braak, C.J.F., Šmilauer, P., 2002a. CANOCO for Windows: software for community ordination (version 4.5). Microcomputer Power, Ithaca New York.

ter Braak, C.J.F., Šmilauer ,P., 2002b. CANOCO Reference Manual and CanoDraw for Windows User's Guide: Software for Canonical Community Ordination (version 4.5). Microcomputer Power, Ithaca, NY.

van Hardenbroek, M., Heiri, O., Parmentier, F.J.W., Bastviken, D., Ilyashuk, B.P., Wiklund, J.A., Hall, R.I., Lottera, A.F., 2013. Evidence for past variations in methane availability in a Siberian thermokarst lake based on $\delta^{13}C$ of chitinous invertebrate remains. Quaternary Science Reviews, 66, 74-84.

Walker, I.R., Levesque, A.J., Cwynar, L.C., Lotter, A.E., 1997. An expanded surface water paleotemperaure inference model for use with fossil midges from Eastern Canada. Journal of Paleolimnology, 18, 165-178.

Walker, I.R., Smol, J.P., Engstrom, D.R., Birks, H.J.H., 1991. An assessment of Chironomidae as quantitative indicators of past climatic change. Canadian Journal of Fisheries and Aquatic Science, 48, 975-987.

Walker, I.R., Cwynar, L.C., 2006. Midges and palaeotemperature reconstruction - the North American experience. Quaternary Science Reviews, 25, 1911-1925.

Wiederholm, T., 1983. Chironomidae of the Holarctic region. Keys and diagnoses. Part 1. Larvae. Entomologica Scandinavica, Suppl. 19.

Woodward, C.A., Shulmeister, J. 2006. New Zealand chironomids as proxies for human-induced and natural environmental change: transfer functions for temperature and lake production (chlorophyll a). Journal of Paleolimnology, 36, 407-429.

Zelentsov, N.I., Shilova, A.I., 1996. Fauna hironomid (Diptera, Chironomidae) Ust-Lenskogo gosudarstvennogo zapovednika. (Fauna of chironomids of Ust-Lenskij natural reserve). Biologia vnutrennih vod (Biology of Inland Waters), 1, 54-61. [In Russian].

Zelentsov, V.V., 2009. The chironomid fauna (Diptera, Chironomidae) of the Khibiny's lakes of Kola Peninsula. Euroasian Entomological Journal, 8, 89–92. [In Russian].

Zolotukhin, V.V., Almukhamedov, A.I., 1988. Traps of the Siberian platform. In: MacDougall, J.D. (Ed.), Continental Flood Basalts. Kluwer Academic Publishers, Dordrecht, pp. 273-310.

Zorina, O.V., 2001. New species of the genera *Cryptotendipes, Dicrotendipes, Microtendipes* and *Stenochironomus* (Diptera, Chironomidae, Chironominae) from the Russian Far East. Vestnik Zoology, 35, 31—38. [In Russian].

Zorina, O.V., 2003. Four new species of non-biting midges of the *Harnischia* complex (Diptera, Chironomidae, Chironominae) from Sakhalin Island (Russian Far East). Eurasian Entomological Journal, 2, 221-225. [In Russian].

Zorina, O.V., 2006a. *Dicrotendipes orientalis* sp.n., a new species of non-biting midges (Diptera, Chironomidae, Chironominae) from the Russian Far East. Russian Entomological Journal, 15, 79-82. [In Russian].

Zorina, O.V., 2006b. Non-biting midges of the genus *Paracladopelma* Harnisch (Diptera: Chironomidae) from the Russian Far East. Zootaxa, 1134, 29-49.

Zorina, O., 2013. A systematic review of the genus *Constempellina* Kieffer (Diptera: Chironomidae) from the Russian Far East, with description of a new species. Zootaxa, 3694, 201–212.

Figure captions

Figure 1. Map of Russia showing the location of the sampled lakes.

Figure 2. Distribution of the lakes along the T July gradient. Abbreviations: WS – West Siberian data set (Self et al., 2011), ES – East Siberian data set (Nazarova et al., 2011), FM – full 268 lake data set (this study), NR – reduced North Russian data set (this study), FE – Far East data set (this study).

Figure 3. Joint chironomid data set ordered by mean July air temperature with the coldest lake at the top and the warmest at the bottom. Chironomids are shown as percentage abundance, for all taxa with abundances greater than 10% at least at one site (268 lakes).

Figure 4. DCA plot of the regions of the data set based on the taxonomic composition.

Figure 5. CCA biplots of (a) 268 Russian lakes and significant environmental variables and (b) common taxa and significant environmental variables; (c) 193 North Russian lakes and significant environmental variables and (d) common taxa and significant environmental variables; (e) 88 Far East lakes and significant environmental variables and (f) common taxa and significant environmental variables. Species fir range 7-100%. Abbreviations: T July: mean July air temperature; WD: water depth; Cond: Log conductivity. Taxon codes correspond with full taxon names listed in ESM 2

Figure 6. Relationship between observed vs estimated T July and the residual (inferred - observed) for (a) the full data set (FM) chironomid-inferred mean July air temperatures 2-component WA-PLS model; (b) the Northern Russia (NR) chironomid-inferred mean July air temperatures 2-component WA-PLS model; (c) the Far East (FE) chironomid-inferred mean July air temperatures 2-component WA-PLS model. Trends in residuals are highlighted with a LOESS smoother (span = 0.45).

Fig 7. T optima of non-rare species in Norwegian (NM, Brooks, Birks, unpublished) and Russian models: WS – West Siberian data set (Self et al., 2011), ES – East Siberian data set (Nazarova et al., 2011), FM – full 268 lake data set (this study), NR – reduced North Russian data set (this study), FE – Far East data set (this study).

Table 1. Summary of environmental data for the 268 lakes.

Parameter	min	max	mean	median	st dev	skew
Latitude, N	53.03	75.39	67.31	67.93	4.86	-0.48
Longitude, E	50.50	163.15	120.06	128.86	29.21	-1.18
Altitude, m a.s.l.	-5	1218	114.6	50	181.0	3.1
pH	4.85	9.96	7.42	7.40	1.05	-0.10
Cond (µS/cm)	2.4	7743.6	245.1	63.0	674.6	7.0
T Jan °C	-45.3	-11.2	-33.9	-36.4	7.6	1.0
T July °C	1.8	18.8	12.1	11.6	4.5	-0.2
CI	21.7	97.0	64.8	65.0	19.4	0.17
Water depth, m	0.1	25	3.3	2.3	3.4	3.3

Table 2. Taxonomic distance (TD) between regions within the data set. Regions are sorted from west to east.

	KOMI	PECHORA	PUTORAN	ANABAR	CY	LENA	INDIGIR	NY	LAPTEV	KOLYMA	KAMCH
KOMI	0.00										
PECHORA	0.98	0.00									
PUTORAN	1.12	0.85	0.00								
ANABAR	1.06	0.68	0.82	0.00							
CY	1.18	1.20	1.22	1.24	0.00						
LENA	1.08	0.74	0.73	0.81	1.18	0.00					
INDIGIR	1.12	0.91	1.07	0.85	1.24	0.97	0.00				
NY	1.32	1.24	1.21	1.30	0.83	1.18	1.27	0.00			
LAPTEV	1.24	1.14	1.22	0.98	1.25	1.21	1.15	1.30	0.00		
KOLYMA	1.11	0.81	0.84	0.78	1.22	0.84	0.84	1.30	1.13	0.00	
KAMCH	1.07	0.97	1.04	0.98	1.24	0.88	0.98	1.35	1.21	0.95	0,00

Table 3. Eigenvalues, cumulative % variance and significance of the CCA axes.

Full data set	Axis 1	Axis 2	Axis 3	Axis 4
Eigenvalues	0.185	0.146	0.046	0.027
Cumulative % variance of taxon data	5.2	9.3	10.6	11.3
Significance (probability) of axis	0.001	0.001	0.001	0.001
Sum of all unconstrained eigenvalues	3.569			
Sum of all canonical eigenvalues	0.404			
NR	Axis 1	Axis 2	Axis 3	Axis 4
Eigenvalues	0.203	0.123	0.052	0.042
Cumulative % variance of taxon data	6.0	9.7	11.2	12.5
Significance (probability) of axis	0.001	0.001	0.001	0.001
Sum of all unconstrained eigenvalues	3.366			
Sum of all canonical eigenvalues	0.465			
Far East	Axis 1	Axis 2	Axis 3	Axis 4
Eigenvalues	0.312	0.154	0.089	0.071
Cumulative % variance of taxon data	9.1	13.6	16.3	18.3
Significance (probability) of axis	0.001	0.001	0.001	0.001
Sum of all unconstrained eigenvalues	3.417			
Sum of all canonical eigenvalues	0.673			

Table 4. Comparison of WA and WA-PLS models for reconstructing mean July air temperature (T July). The best models are given in bold

	N lakes in the model	N taxa in the model	Model		r^2_{jack}	RMSEP (% RMSEP reduced from the previous component)	Max bias $_{Jack}$
Full data set model	268	176	WA	Inverse	0.69	2.50	3.43
				Classical	0.69	2.77	2.85
			WA-PLS	1 component	0.69	2.50	3.50
				2 component	**0.73**	**2.34 (6.47)**	**2.52**
				3 component	0.71	-3.39	2.51
				4 component	0.69	-4.05	2.55
				5 component	0.66	-7.25	2.28
North Russia reduced model	193	163	WA	Inverse	0.81	1.64	3.04
				Classical	0.81	1.70	2.79
			WA-PLS	1 component	0.80	1.65	2.98
				2 component	**0.87**	**1.35 (18.3)**	**2.21**
				3 component	0.87	1.32	2.27
				4 component	0.87	1.34	2.49
				5 component	0.85	1.43	2.67
Far East reduced model	88	135	WA	Inverse	0.76	1.45	4.02
				Classical	0.76	1.48	4.16
			WA-PLS	1 component	0.75	1.46	4.03
				2 component	**0.81**	**1.30 (11.4)**	**3.39**
				3 component	0.81	1.27	3.78
				4 component	0.81	1.30	3.86
				5 component	0.79	1.34	3.86

Table 5. Number of the taxa with more than 10 occurrences and relationship of the taxa in the full data set (FM), the reduced Northern Russia data set (NR) and the reduced Far East data set (FE) to T July.

	FM	NR	FE
n taxa with more than 10 occurrences	94	89	49
Highly significant response to T July, n taxa (%)	60(63.8%)	43(48.4%)	23 (47.0%)
Significant response to T July, n taxa (%)	21(22.3%)	25 (28.0%)	16 (32.6%)
No response, n taxa (%)	13(13.9%)	21 (23.6%)	10 (20.4%)
Significant/ not significant, %	86.1/13.9	76.5/23.5	79.6/20.4

Figure 1

Click here to download high resolution image

Figure 2

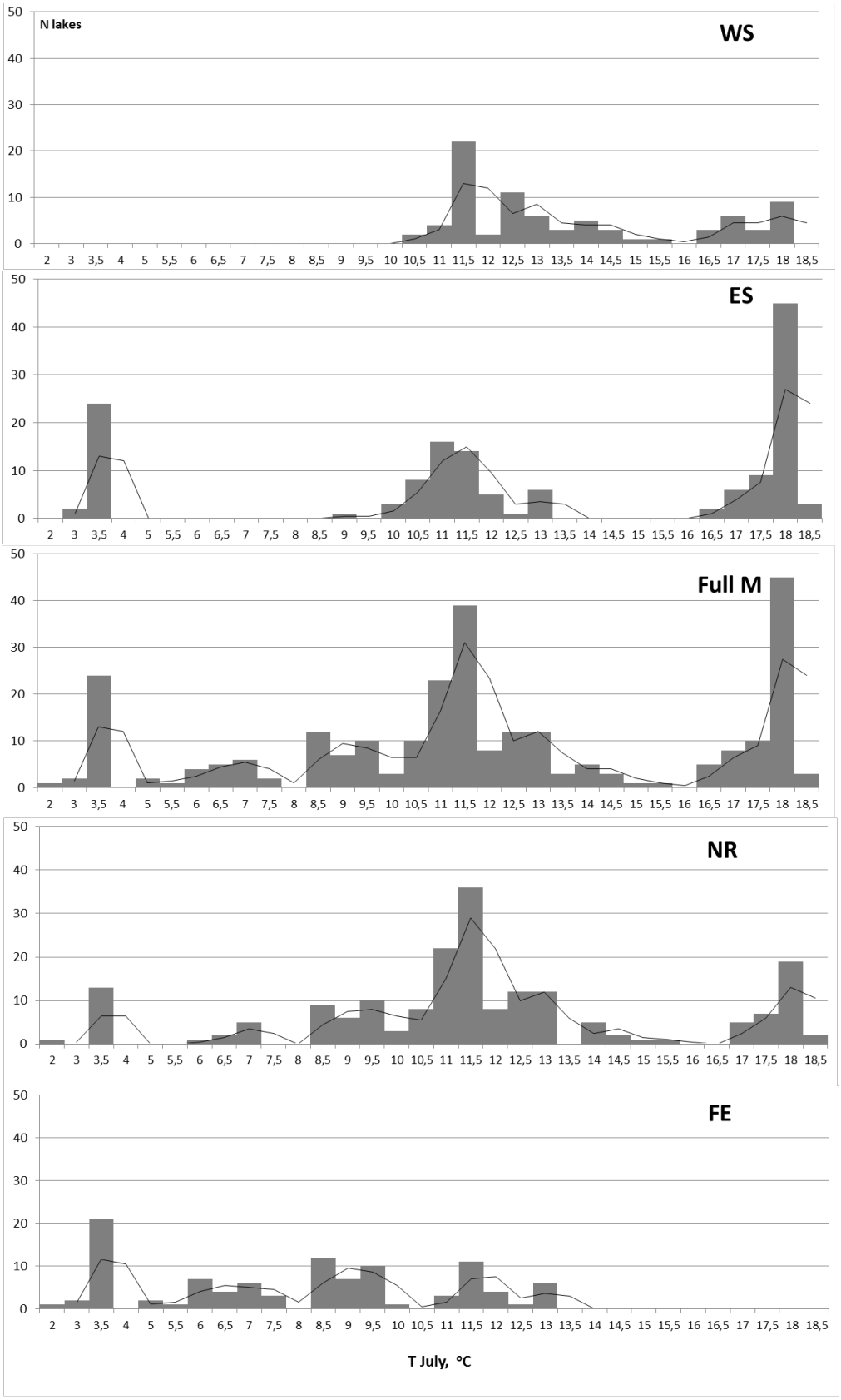


Fig.2. Distribution of the lakes along the T July gradient. Abbreviations: WS – West Siberian data set, ES – East Siberian data set, Full M – full set of the data, NR – reduced combined North Russian data set. FE – Far East data set

Figure 3
Click here to download high resolution image

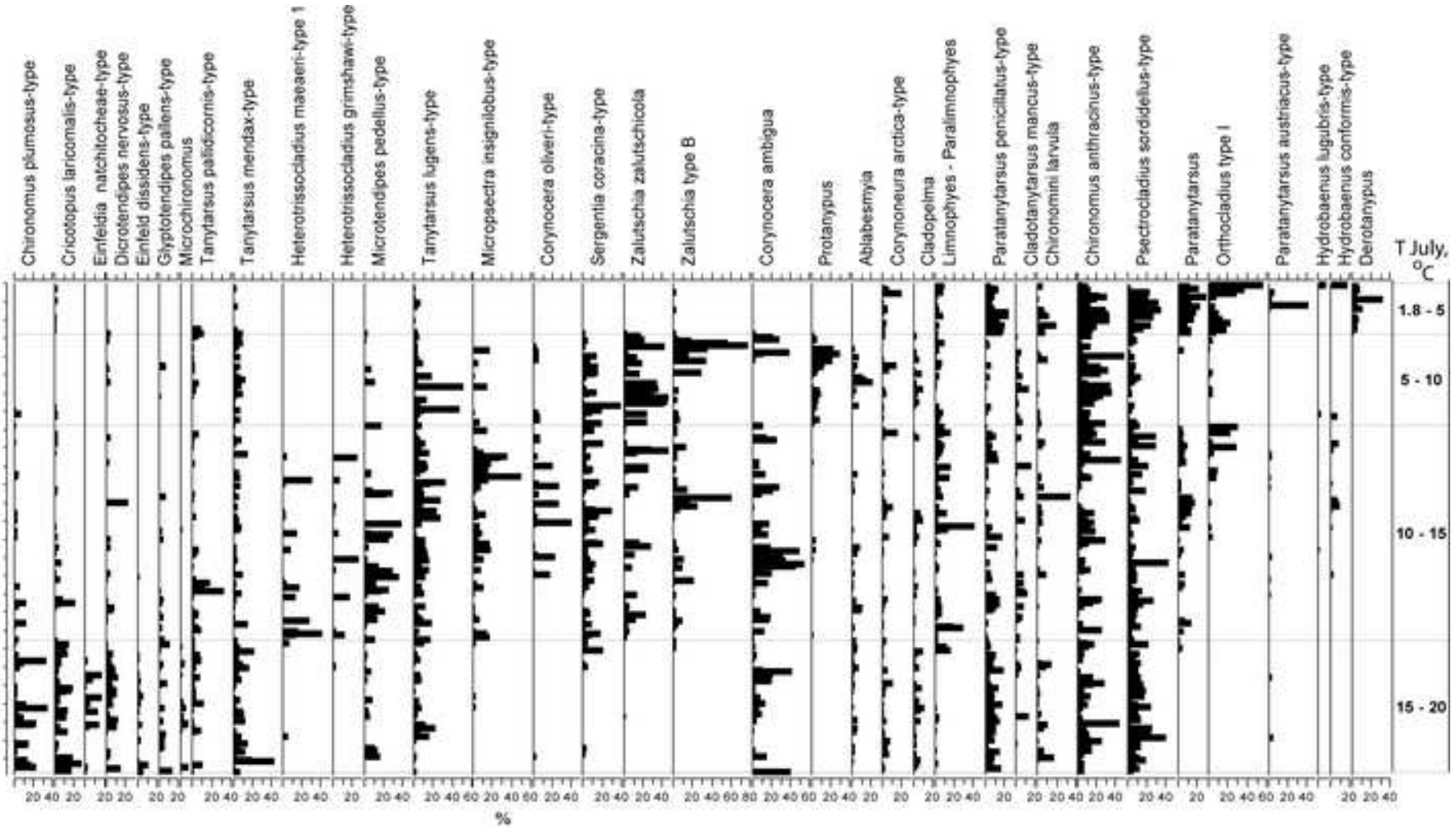


Figure 4

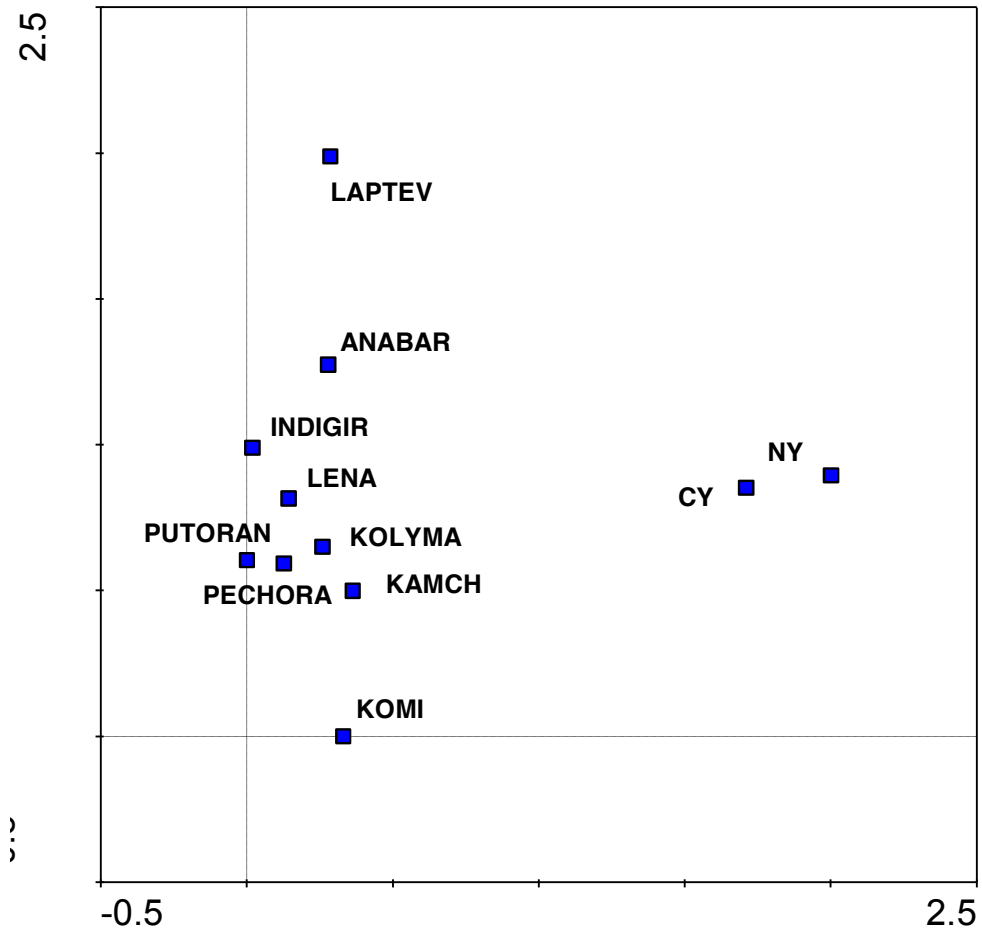


Figure 4. DCA of the regions of the dataset based on the taxonomic composition

Figure 5 ab

Click here to download high resolution image

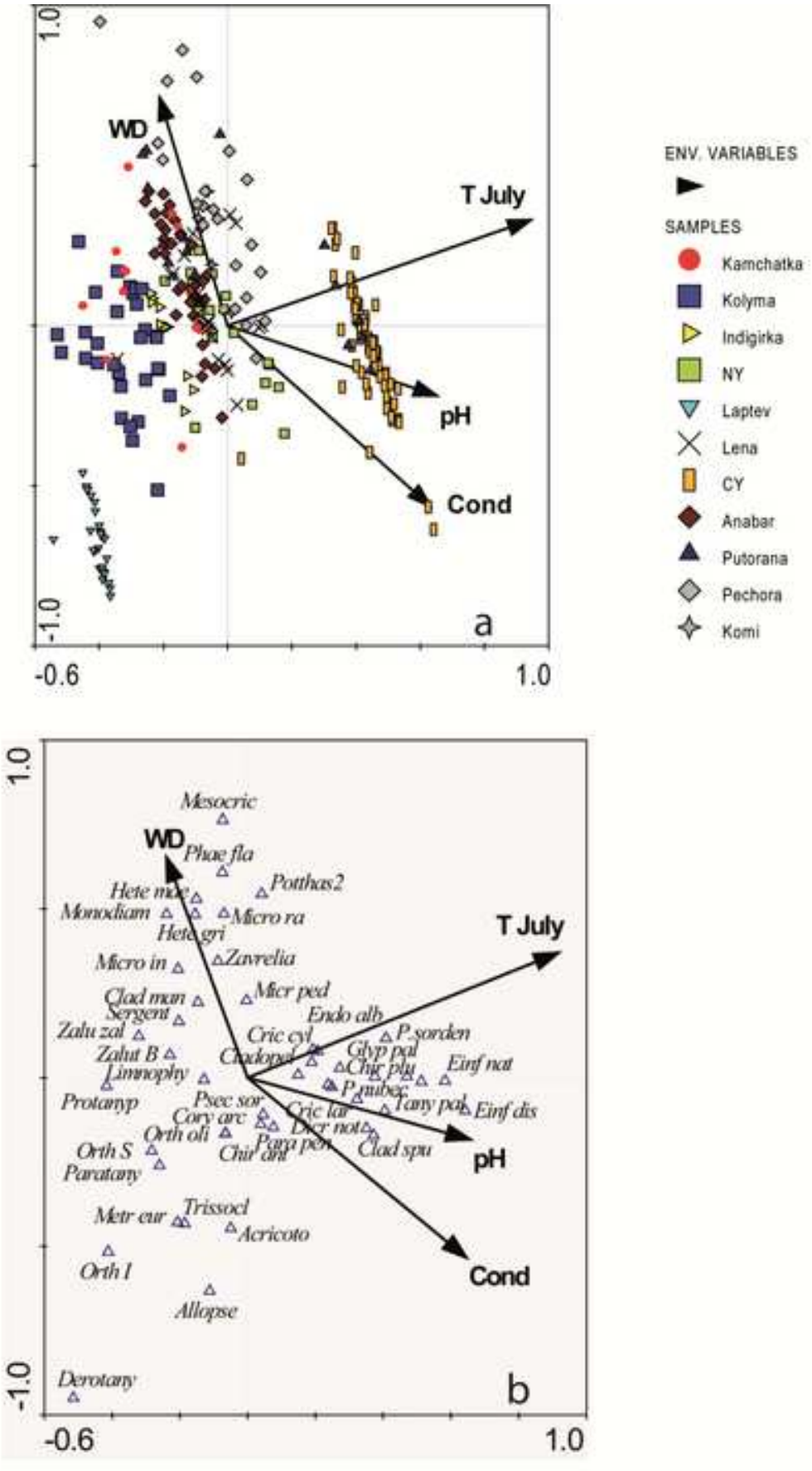
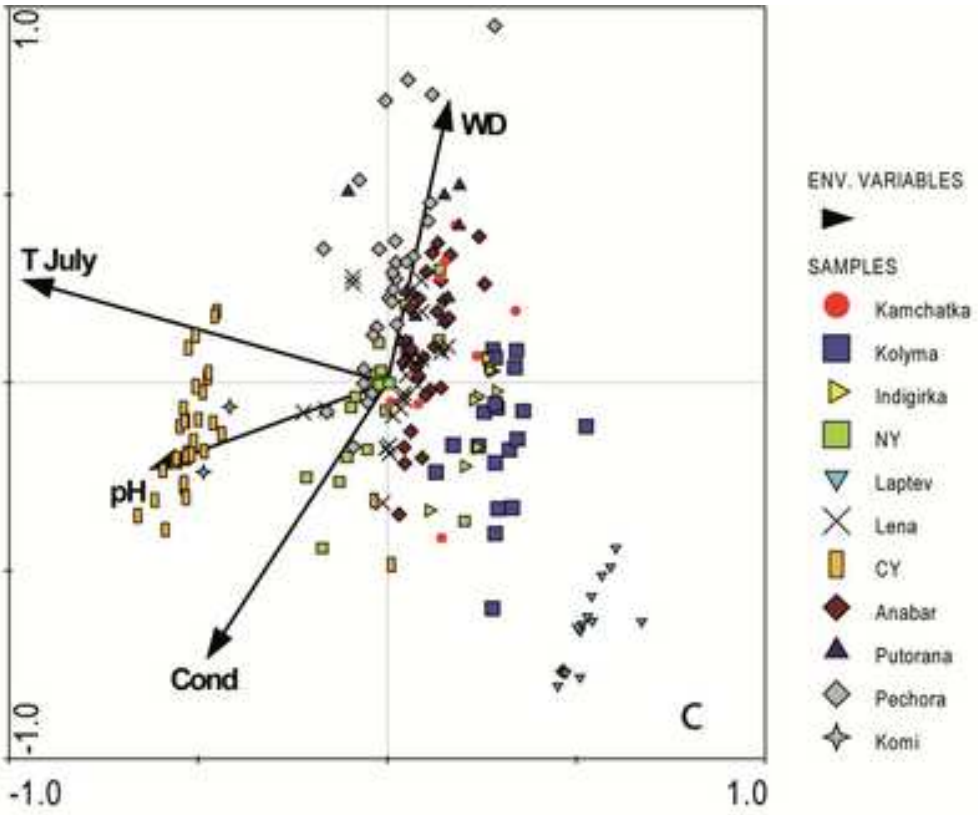


Figure 5 cd

Click here to download high resolution image



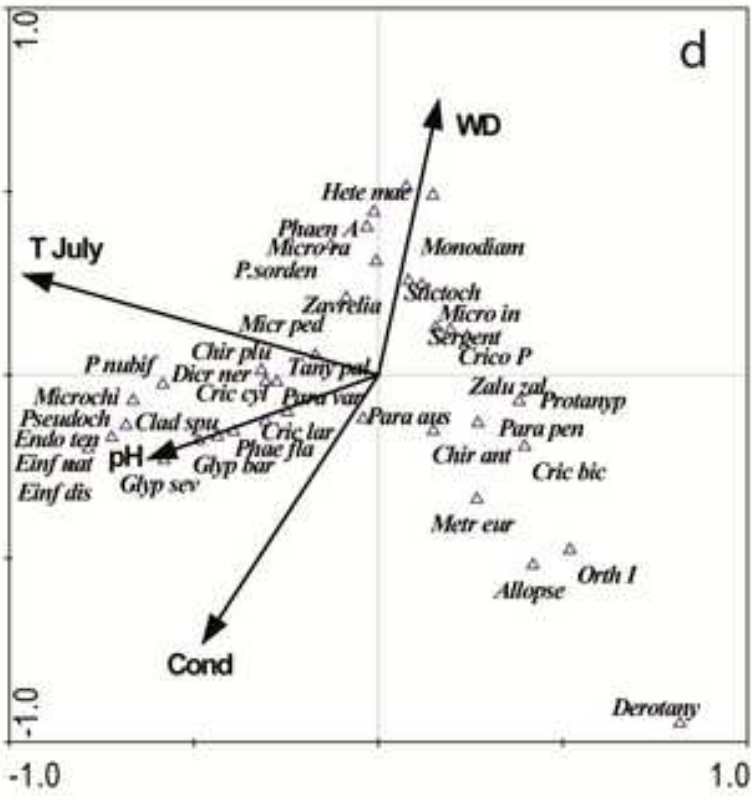
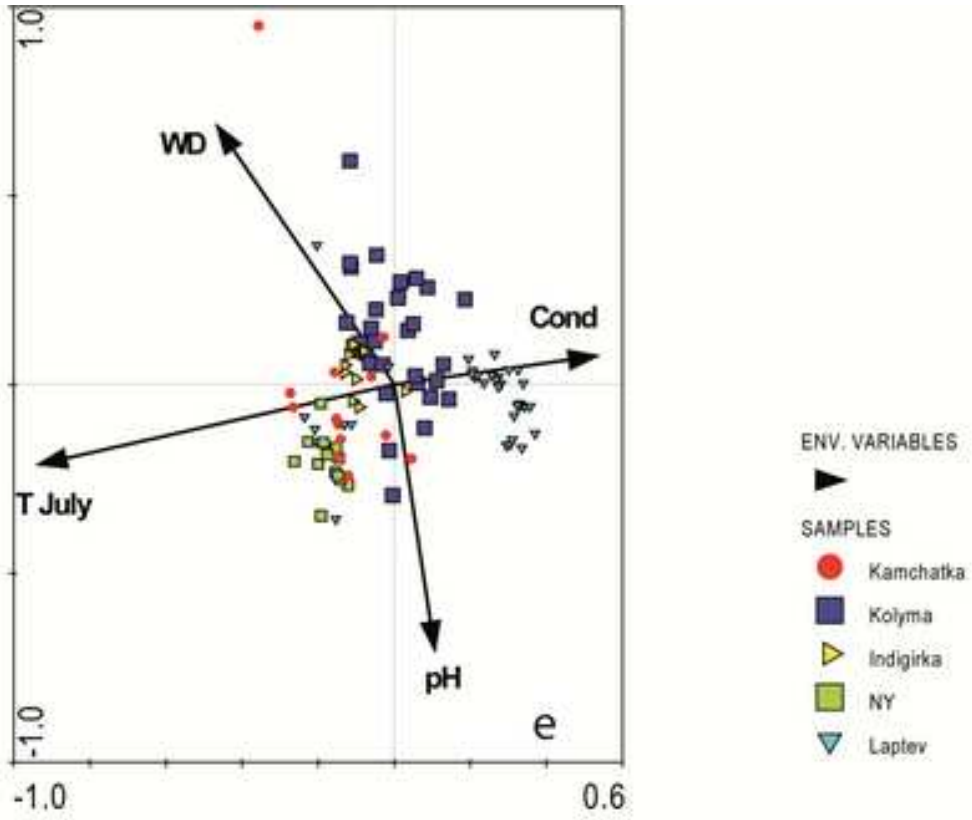


Figure 5 ef

Click here to download high resolution image



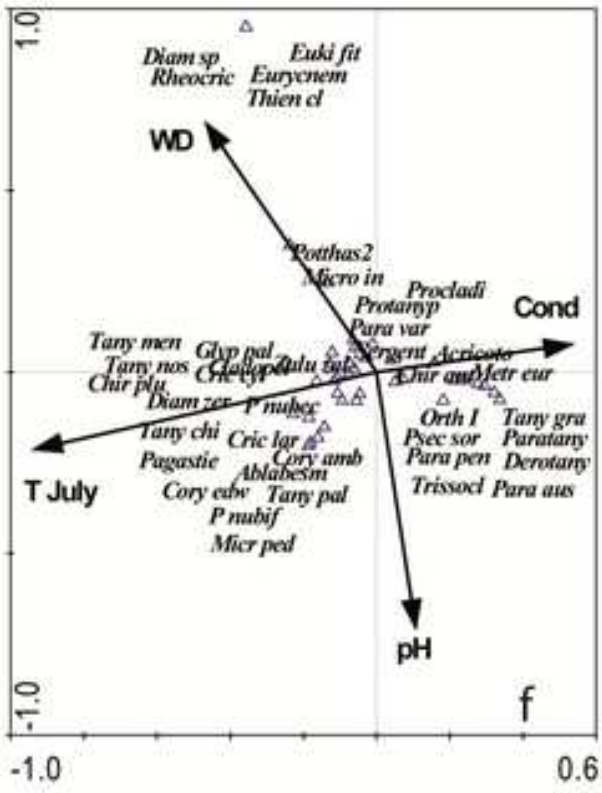


Figure 6
Click here to download high resolution image

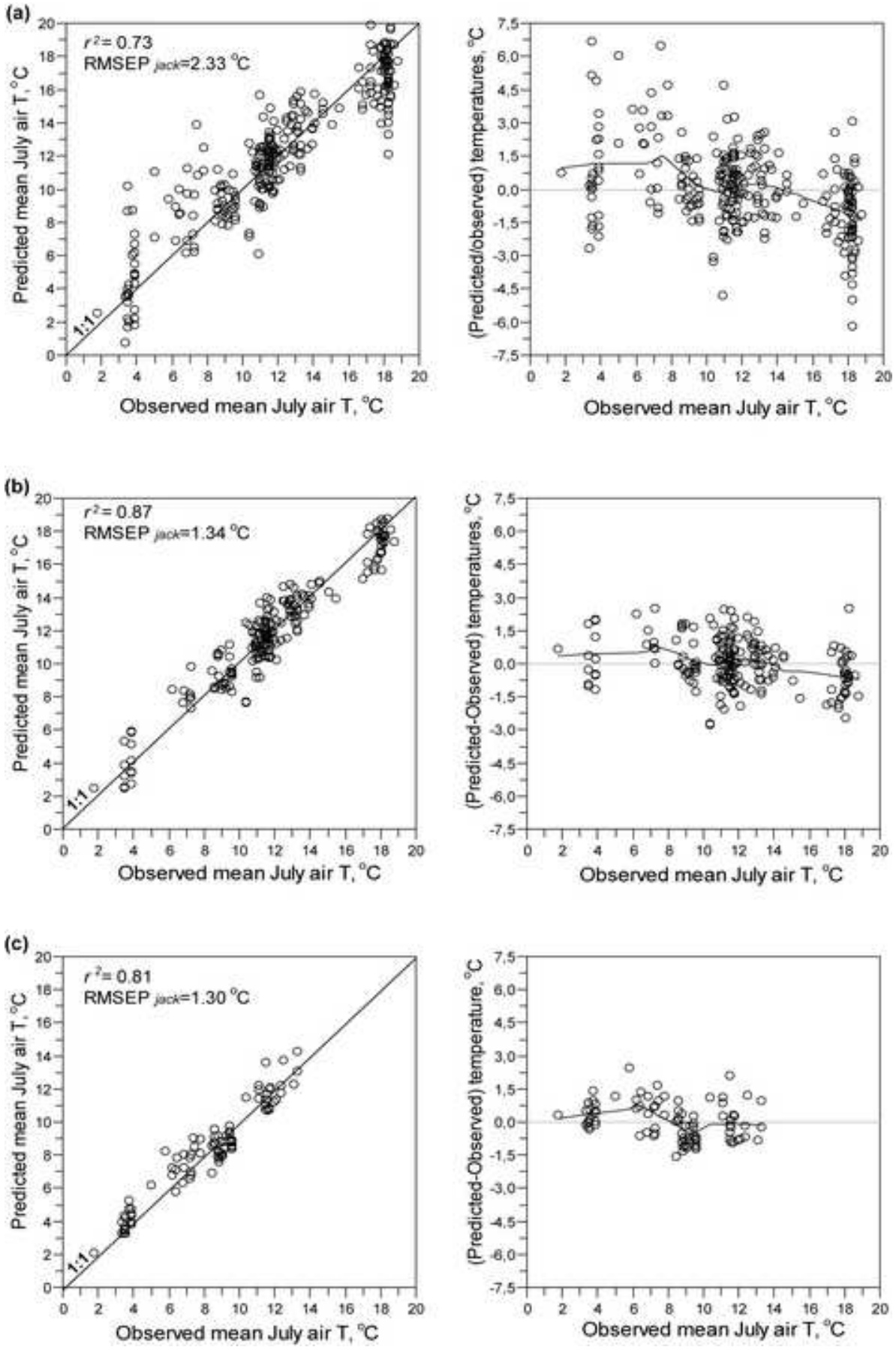
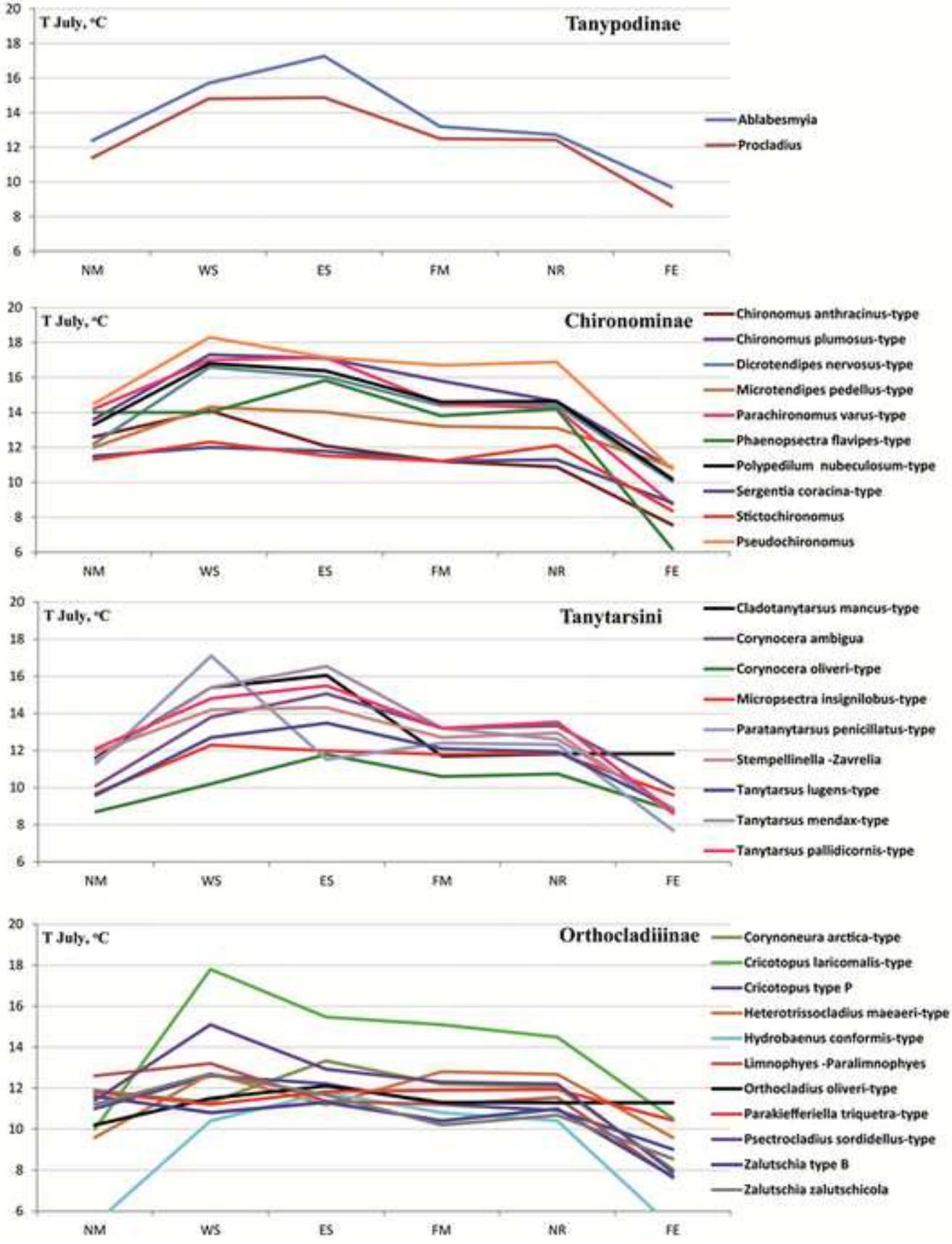


Figure 7
Click here to download high resolution image



ESM 1

Electonic supplementary material 1

Environmental variables, canonical coefficients and t-values of significant environmental variables used in the CCA

Full Data set	Canonical coefficients				t-values				Interset correlation			
	Axis 1	Axis 2	Axis 3	Axis 4	Axis 1	Axis 2	Axis 3	Axis 4	Axis 1	Axis 2	Axis 3	Axis 4
T_{July}	0.86	0.64	0.04	0.64	16.49	10.21	0.39	5.59	0.78	0.26	-0.02	-0.02
EC	0.22	-0.70	-1.01	0.01	4.14	-11.1	-10.80	0.05	0.52	-0.44	-0.30	-0.14
WD	-0.23	0.53	-0.59	-0.65	-5.20	10.04	-7.51	-6.72	-0.18	0.55	-0.34	-0.22
pH	0.01	-0.09	0.74	-1.15	0.11	-1.26	7.17	-9.16	0.54	-0.17	0.18	-0.37
NR Data set	Canonical coefficients				t-values				Interset correlation			
	Axis 1	Axis 2	Axis 3	Axis 4	Axis 1	Axis 2	Axis 3	Axis 4	Axis 1	Axis 2	Axis 3	Axis 4
T_{July}	-0.86	0.46	-0.34	0.49	-17.77	5.53	-2.86	4.17	-0.83	0.19	-0.05	0.03
EC	-0.14	-0.66	-0.81	-0.50	-2.91	-7.76	-6.69	-4.20	-0.41	-0.51	-0.19	-0.22
WD	0.13	0.52	-0.35	-0.85	2.80	6.70	-3.16	-7.79	0.14	0.53	-0.15	-0.34
pH	-0.15	-0.07	1.05	-0.59	-2.93	-0.77	8.32	-4.74	-0.54	-0.16	0.34	-0.27
FE Data set	Canonical coefficients				t-values				Interset correlation			
	Axis 1	Axis 2	Axis 3	Axis 4	Axis 1	Axis 2	Axis 3	Axis 4	Axis 1	Axis 2	Axis 3	Axis 4
T_{July}	-0.83	-0.28	-0.72	0.05	-12.84	-2.35	-7.77	0.48	-0.81	-0.15	-0.17	0.15
EC	0.26	0.25	-0.92	0.49	4.21	2.10	-10.24	4.39	0.46	0.05	-0.46	0.43
WD	-0.16	0.70	0.55	0.59	-2.53	6.17	6.20	5.37	-0.41	0.47	0.24	0.34
pH	0.13	-0.63	0.65	0.61	2.07	-5.40	7.27	5.47	0.09	-0.48	0.16	0.48

Electronic supplementary material 2

List of taxon codes and taxon names.

Taxon code	Taxon name
Abiskomy	Abiskomyia
Ablabesm	Ablabesmyia
Acalcare	Acalcarella
Acamptoc	Acamptocladius
Brillia	Brillia
Chaetocl	Chaetocladius
Chir ant	Chironomus anthracinus
Chir lar	Chironomini larvula
Chir plu	Chironomus plumosus
Clad man	Cladotanytarsus mancus
Clad spu	Cladotanytarsus mancus 'with spur'
Cladopel	Cladopelma
Constemp	Constempellina - Thienemanniola
Cory amb	Corynocera ambigua
Cory arc	Corynoneura arctica
Cory edw	Corynoneura edwardsi
Cory lob	Corynoneura lobata
Cory oli	Corynocera oliveri
Cric cyl	Cricotopus cylindraceus
Cric int	Cricotopus intersectus
Cric lar	Cricotopus laricomalis
Cric syl	Cricotopus sylvestris
Cric tri	Cricotopus trifaciatus
Crico C	Cricotopus type C
Crico ob	Cricotopus obnixus
Crico P	Cricotopus type P
Cryptoch	Cryptochironomus
Dicroten	Dicrotendipes
Diplocla	Diplocladius
Einfeldia	Einfeldia
Endo alb	Endochironomus albipennis
Endo imp	Endochironomus impar
Endo ten	Endochironomus tendens
Euki cla	Eukiefferiella claripennis
Euki dev	Eukiefferiella devonica
Georthoc	Georthocladius
Glyp bar	Glyptotendipes barbipes
Glyp pal	Glyptotendipes pallens
Glyp sev	Glyptotendipes severini
Hete api	Heterotanytarsus apicalis
Hete bru	Heterotrissocladius maeaeri
Hete gri	Heterotrissocladius grimshawi
Hete mar	Heterotrissocladius marcidus
Hete sub	Heterotrissocladius subpilosus

Hydr con	Hydrobaenus conformis
Hydr joh	Hydrobaenus johannseni
Hydr lug	Hydrobaenus lugubris
Lasiodia	Lasiodiamesa
Lauterbo	Lauterborniella
Limnophy	Limnophyes - Paralimnophyes
Lipinell	Lipiniella
Macropel	Macropelopia
Mesocric	Mesocricotopus
Mesopsec	Mesopsectrocladius
Metr eur	Metriocnemus eurynotus
Metr fus	Metriocnemus fuscipes
Micr ped	Microtendipes pedellus
Micro in	Micropsectra insignilobus
Micro ra	Micropsectra radialis
Micro ry	Microtendipes rydalensis
Microchi	Microchironomus
Monodiam	Monodiamesa
Nano bra	Nanocladius branchicolus
Nano rec	Nanocladius rectinervis
Orth oli	Orthocladius oliveri
Orth tri	Orthocladius trigonolabis
Orth und	Orthocladiinae undiff.
Ortho I	Orthocladius type I
Ortho S	Orthocladius type S
Orthocla	Orthocladius/Cricotopus
Pagastie	Pagastiella
Para aus	Paratanytarsus austriacus
Para bat	Parakiefferiella bathophila
Para nig	Parakiefferiella nigra
Para nud	Paratendipes nudisquama
Para pen	Paratanytarsus penicillatus
Para tri	Parakiefferiella triquetra
Para var	Parachironomus varus
Parachae	Parachaetocladius
Paraclad	Paracladius
Paraclop	Paracladopelma
Parak A	Parakiefferiella type A
Paralaut	Paralauterborniella
Parameri	Paramerina
Paraphae	Paraphaenocladius - Parametriocnemus
Paratany	Paratanytarsus undiff.
Phae fla	Phaenopsectra flavipes
Phaen A	Phaenopsectra type A
Pogonocl	Orthocladius (Pogonocladius) consobrinus
Poly nub	Polypedilum nubeculosum
Poly sor	Polypedilum sordens
Procladi	Procladius
Prop lac	Propsilocerus lacustris

Prop yak	Propsilocerus jacuticus
Protanyp	Protanypus
Psec psi	Psectrocladius psilopterus
Psec sep	Psectrocladius septentrionalis
Psec sor	Psectrocladius sordidellus
Pseudoch	Pseudochironomus
Pseudort	Pseudorthocladius
Pseudosm	Pseudosmittia
Rheo cha	Rheocricotopus chalybeatus
Rheo eff	Rheocricotopus effusus
Rheotany	Rheotanytarsus
Serg cor	Sergentia coracina
Smittia	Smittia - Parasmittia
Stempell	Stempellina
Stenochi	Stenochironomus
Stictoch	Stictochironomus
Synortho	Synorthocladius
Tany chi	Tanytarsus chinyensis
Tany gla	Tanytarsus glabrescens
Tany lug	Tanytarsus lugens
Tany men	Tanytarsus mendax
Tany nos	Tanytarsus 'no spur'
Tany pal	Tanytarsus pallidicornis
Tanypus	Tanypus
Thie cla	Thienemanniella clavicornis
Thienmia	Thienemannimyia-group
Trissocl	Trissocladius
Unid Cor	Corynoneura undiff.
Unid Pen	Pentaneurini undiff.
Unid pod	Indeterminate Tanypodinae
Unid Tan	Tanytarsini undiff.
Xeno	Xenochironomus
Zalu muc	Zalutschia mucronata
Zalu zal	Zalutschia zalutschicola
Zalut A	Zalutschia type A
Zalut B	Zalutschia type B
Zavreie	Zavreliella
Zavrelia	Stempellinella - Zavrelia

Electronic supplementary material 3

Number of occurrences (N), maximum abundance (Max), N2, response to T July (WS model: HOF model; ES, FM, NR and FE models: Resp, value for significance of the relationship with the T July based on generalized linear response model, set to a quadratic degree and Poisson distribution p<0,05: - not significant, significant, small x; highly significant, p<0,001 capital X), and WA optima (Opt) and tolerances (tol) for taxa with more than 10 occurrences in data-sets. HOF model I shows no response to July air temperature; II a sigmoidal increasing or decreasing response, III a response which reaches a plateau; IV a unimodal response and V a skewed unimodal response. Data sets: WS – West Siberian, ES – East Siberian, FM – full model, NR – North Russian, FE– Far East.

	WS					ES						FM						NR						FE					
	N	Max	N2	Opt	HOF	N	Max	N2	Opt	tol	Resp	N	Max	N2	Opt	tol	Resp	N	Max	N2	Opt	tol	Resp	N	Max	N2	Opt	tol	Resp
Tanypodinae																													
Ablabesmyia	46	8.2	32.5	15.7	V	16	2.4	15.3	17.2	0.7	X	99	4.7	82.6	13.2	3.9		73	3.9	61.4	12.7	3.2	-	36	4.7	28.0	9.7	2.2	X
Procladius	77	13.6	50.3	14.8	IV	14	3.0	12.0	14.9	3.8	X	188	3.9	166.0	12.5	4.1	x	143	3.7	126.7	12.4	3.2	X	60	3.0	52.7	8.6	2.2	X
Chironomini																													
Chironomus anthracinus-type	66	26.7	29.2	14.1	II	126	6.8	104.1	12.1	5.6	X	220	7.1	177.1	11.2	4.9	X	158	7.1	124.3	10.9	4.1	X	76	7.1	63.4	7.6	3.0	x
Chironomus plumosus-type	49	39.4	21.2	17.3	II	40	6.0	29.6	17.1	1.7	X	76	6.0	58.0	15.8	3.2	X	58	4.6	47.7	14.6	3.3	X	18	3.6	15.0	10.8	1.3	X
Cladopelma	50	11.2	27.3	15.2	II	54	3.4	47.5	16.4	2.7	X	110	3.4	94.6	13.8	4.0	X	80	3.3	68.5	12.9	3.6	-	34	3.3	30.0	9.3	1.8	X
Dicrotendipes nervosus-type	60	23.6	33.0	16.6	II	56	3.9	48.3	16.1	2.5	X	107	4.9	88.0	14.4	3.8	x	82	4.9	65.8	14.4	3.3	X	31	3.0	26.6	10.0	2.3	X
Microtendipes pedellus-type	43	27.2	20.3	14.3	IV	37	6.4	27.4	14.0	3.0	X	82	6.4	60.7	13.2	3.1	X	66	6.4	48.3	13.1	2.5	X	29	6.4	20.9	10.9	1.7	-
Parachironomus varus-type	28	6.6	18.3	17.0	II	38	2.8	34.4	17.2	1.8	X	66	2.9	57.8	14.5	4.4	X	43	2.5	38.7	14.2	3.8	X	20	2.9	17.0	8.8	2.5	X
Phaenopsectra flavipes-type	15	2.8	12.2	14.0	IV	14	3.0	12.0	15.8	2.6	x	16	2.0	14.8	13.8	2.4	X	15	2.0	13.8	14.2	1.6	X	1	1.1	1.0	6.2	2.0	-
Polypedilum nubeculosum-type	52	8.3	36.5	16.8	IV	54	4.5	46.7	16.4	2.5	X	77	4.5	67.0	14.6	4.1	X	49	4.5	42.2	14.6	3.7	X	31	4.5	26.4	10.2	2.0	X
Sergentia coracina-type	50	30.8	25.2	12.0	II	34	6.3	26.5	11.8	2.2	X	127	6.3	101.5	11.2	2.7	X	108	6.3	86.5	11.3	2.1	X	47	6.3	36.5	8.8	1.7	X
Stictochironomus	29	8.3	17.7	12.3	II	12	3.7	10.2	11.5	3.6	x	48	3.7	40.5	11.2	3.1	X	42	3.7	35.7	12.1	2.1	X	8	3.7	6.7	8.4	2.2	-
Pseudochironomus	18	9.5	12.7	18.3	II	24	3.1	21.0	17.2	1.4	X	30	3.1	25.6	16.7	3.0	X	23	3.1	19.4	16.9	2.4	X	7	2.0	5.9	10.8	2.7	-
Tanytarsini																													
Cladotanytarsus mancus-type	71	16.7	40.1	15.4	II	5	3.9	3.4	16.1	3.2	X	80	4.1	65.7	11.7	3.1	X	69	4.1	55.8	11.8	2.7	X	69	4.1	55.8	11.8	2.7	X
Corynocera ambigua	60	56.5	23.9	13.8	IV	48	6.8	37.4	15.1	2.9	X	102	7.5	74.0	13.2	3.6	X	86	7.5	63.0	13.3	3.2	-	28	6.8	21.6	10.0	2.5	X
Corynocera oliveri-type	20	41.8	8.5	10.2	IV	17	2.5	15.5	11.8	4.1	-	57	6.5	40.0	10.6	2.7	X	47	6.5	32.1	10.7	1.8	X	21	2.9	18.2	8.8	2.2	X
Micropsectra insignilobus-type	39	20.1	18.2	12.3	IV	38	6.0	29.3	12.0	2.3	X	89	7.1	65.7	11.8	2.1	X	79	7.1	59.4	11.9	1.8	X	16	7.1	9.7	9.6	1.4	X
Micropsectra radialis-type	14	18.2	4.8	12.8	IV	7	2.1	6.4	14.7	3.1	-	24	4.3	19.2	13.2	3.2	X	24	4.3	19.2	13.2	3.2	-	2	1.7	2.0	10.4	2.7	-
Paratanytarsus penicillatus-type	64	24.1	36.7	*17.1*	II	118	6.8	97.7	11.5	5.9	X	169	5.1	139.8	12.4	5.4	X	120	4.6	97.8	12.3	4.6	X	44	5.1	36.6	7.7	3.8	X
Stempellinella -Zavrelia	41	9.4	26.2	14.2	IV	8	2.6	7.4	14.3	3.6	-	57	3.1	49.5	12.7	3.3	X	46	3.1	40.2	12.9	2.6	X	12	2.3	10.7	8.6	2.2	X
Tanytarsus lugens-type	75	28.8	39.6	12.7	V	80	5.7	65.2	13.5	4.0	X	175	7.2	141.3	12.1	3.6	X	141	7.2	113.5	11.9	3.0	X	54	7.2	41.4	8.9	2.2	X
Tanytarsus mendax-type	63	22.5	35.8	15.4	II	35	3.6	30.8	16.5	2.8	X	147	6.7	122.6	13.2	4.4	X	103	4.1	88.3	12.6	3.5	X	53	3.7	47.2	8.9	2.3	X
Tanytarsus pallidicornis-type	58	33.3	29.5	14.8	IV	17	2.9	15.0	15.5	3.0	X	101	5.8	85.1	13.2	4.0	x	78	5.8	64.8	13.5	3.2	x	26	2.8	23.6	8.7	2.3	X
Orthocladiinae																													
Corynoneura arctica-type	31	11.3	17.4	11.3	II	72	4.6	60.7	13.3	5.5	x	129	4.6	109.2	12.2	4.9	x	90	3.5	78.9	12.1	3.9	-	37	3.9	32.0	8.0	3.1	-
Cricotopus laricomalis-type	54	11.3	26.4	17.8	II	86	5.4	70.5	15.5	3.7	X	116	5.4	92.4	15.1	4.0	X	84	4.8	67.3	14.5	3.7	X	19	4.8	15.5	10.5	3.0	X
Cricotopus type P	24	24.4	15.2	12.6	IV	20	3.0	16.7	12.2	3.0	X	63	3.0	54.6	11.2	2.7	-	54	3.0	46.5	10.9	2.1	X	24	3.0	21.3	9.0	1.9	X
Heterotrissocladius grimshawi-type	13	6.2	5.8	13.6	IV	7	5.2	5.0	11.5	0.5	X	20	5.2	14.6	12.7	2.0	-	17	5.2	12.8	12.9	1.9	X						-
Heterotrissocladius maeaeri-type	25	17.9	9.8	12.7	IV	7	5.7	4.5	11.2	0.5	X	33	6.5	22.2	12.8	1.7	X	31	6.5	20.4	12.7	1.4	X	1	1.0	1.0	9.6	2.0	-
Hydrobaenus conformis-type	17	42.2	8.4	10.4	II	10	3.1	8.4	11.8	4.0	-	27	4.4	20.8	10.8	3.9	X	25	4.4	19.3	10.4	3.6	X	4	4.4	3.1	5.2	4.5	-
Limnophyes -Paralimnophyes	55	30.4	17.6	13.2	II	75	6.5	59.8	11.4	4.7	X	155	6.5	123.7	11.2	4.1	-	116	6.5	90.9	11.5	3.0	X	45	3.7	39.4	7.6	2.9	-
Orthocladius oliveri-type	29	11.0	17.1	11.5	II	35	3.2	29.8	12.1	6.1	x	89	3.9	75.3	11.3	4.5	X	65	3.9	54.0	11.3	3.4	x	65	3.9	54.0	11.3	3.4	-

Parakiefferiella bathophila-type	40	7.4	27.3	14.5	I	38	3.2	32.8	14.9	3.6	x	74	3.2	65.4	13.7	3.6	x	55	2.6	50.1	13.0	2.5	-	10	2.1	8.9	9.5	2.5	-
Parakiefferiella triquetra-type	16	7.4	10.4	11.2	II	5	1.5	4.7	11.8	0.8	-	23	2.7	20.2	11.9	1.3	x	23	2.7	20.2	11.9	1.3	X	3	1.8	2.8	10.4	2.0	-
Paraphaenocladius	12	4.9	8.3	15.6	IV	9	2.4	8.3	14.4	3.5	-	33	2.8	30.0	14.0	4.2	-	23	2.8	20.5	13.2	3.8	-	4	1.5	4.0	8.5	4.1	X
Psectrocladius septentrionalis-type	21	5.7	11.0	13.5	IV	5	2.5	4.7	10.2	3.3	X	30	3.4	24.1	11.8	2.7	X	25	2.8	20.3	12.2	1.5	-	6	3.4	5.6	9.8	1.7	-
Psectrocladius sordidellus-type	87	27.0	57.2	15.1	I	132	6.6	110.4	12.9	5.4	X	228	6.6	191.8	12.3	4.9	X	163	6.6	138.0	12.2	3.9	X	65	6.6	54.2	7.9	3.3	X
Pseudosmittia	13	3.8	9.9	14.5	IV	13	1.9	12.0	10.2	5.2	x	30	2.8	26.7	10.7	5.0	X	20	1.9	18.4	11.0	3.8	-	9	2.8	7.6	4.9	2.7	X
Zalutschia type B	47	61.3	10.9	10.8	II	25	4.0	21.7	11.3	2.4	X	99	8.9	66.7	10.4	2.9	X	83	7.8	59.6	11.0	2.4	X	33	8.9	21.0	7.7	1.9	-
Zalutschia zalutschicola	24	24.3	10.5	12.7	IV	32	7.0	21.5	11.7	1.4	X	90	7.0	63.2	10.2	2.5	X	77	7.0	54.7	10.7	2.2	X	40	7.0	29.8	8.6	1.6	X